Working Paper Series

HEARIN CENTER FOR ENTERPRISE SCIENCE

HCES-02-03

Comparing Sequential and Joint Mesh Network Capacity Optimization Using An Experimental Design Approach

by

Mark W. Lewis

The University of Mississippi

Director, Keith Womer
School of Business Administration
The University of Mississippi
Post Office Box 1848
University, MS 38677-1848
(662) 915-5820
http://hces.bus.olemiss.edu

Comparing Sequential and Joint Mesh Network Capacity Optimization Using An Experimental Design Approach

Mark W. Lewis

School of Business, University of Mississippi, University, MS 38677, USA. mlewis@bus.olemiss.edu

Abstract — In a robust telecommunications network, additional capacity must be allocated to account for restoration routing after the loss of a network component. Joint capacity optimization takes into account the relationship between working paths and restoration paths in order to generate robust networks with less total capacity than when the capacity for working and restoration paths are optimized sequentially. This paper uses experimental design methodology to show that joint optimization always yielded much better capacity solutions (> 25% improvement) and that changing network characteristics had little effect on the per cent improvement of joint over sequential optimization. However, as network connectivity and edge capacities increase, total capacity requirements decrease for both methods and subsequently the total capacity benefit of JCA over SCA decreases. The experimental design approach is used to mathematically model the effects of, and interactions between, demand distribution, average node degree, fiber-pairs per edge, and algorithm type. The model is shown to accurately predict, within an average of 5%, the total capacity needed at intermediate demand levels.

Keywords : Capacity Placement; Integer Programming, Experimental Design

I. Introduction

In a telecommunications network, capacity is measured in units of bandwidth and the demands between origin-destination pairs represent multiple commodities. Network problems are often modeled as capacitated multi-commodity flow network problems [15] with integer side constraints, which are known to be hard problems [19]. Because of the large amount of information being transported over networks, the physical loss of a network component implies the loss of large amounts of data. Re-routing around the failed component to restore the network requires implementing restoration routes using spare capacity. This investigation compares two routing optimizations (sequential and joint) that minimize capacity requirements for a network with a single edge failure. Proving a routing and capacity allocation solution is optimal may take weeks of computer time and large amounts of computer memory, thus, an efficient testing method is needed. Experimental design techniques yield insightful results with a minimum of testing.

Managers of a network are interested in how various parameters, some under their control, some not, affect their network. For example, they need to be able to appropriately allocate capacity to handle changes in demand and they need to be able to ascertain the effects of physical changes to their network as it is managed over time. To this end, this investigation studies a generic nation-wide transport network topology under changing: algorithm type; node degree; demand pattern; and modular edge capacity limits. While the conclusions made might be generalized to all networks, this research focuses on the application of experimental design (DOE) techniques to gain specific network insights and guide further investigations.

In this paper, *mesh* networks with diverse working/restoration routing are studied, as opposed to robust networks designed as stars or rings. Related ring routing and capacity allocation problems are studied in [3, 4, 11, 20]. Telecommunication network connections between nodes are defined as bidirectional *edges*, where each edge is composed of a group of fiber optic strands (or fiber-pairs). Other commonly used terms for an edge are span, arc, and link. The fiber-pairs are terminated by circuit cards contained in equipment such as a digital cross connect that is housed in the node. The number and type of circuit cards is dependent on equipment characteristics (visit the home pages of Cisco, Nortel, Alcatel, Lucent or other telecommunication equipment manufacturer). Basically, only certain *modular* capacities can be installed on an edge, e.g. combinations of OC-12, OC-48, and OC-192. In the model and results presented here, the term *slots*, for circuit card slots, is used to describe modular edge capacity limits.

The two basic mesh restoration categories are edge restoration [1, 8, 12], which re-routes around the end nodes of the failed edge and *path restoration* [13, 16] which re-routes between the affected origins and destinations. Path based restoration is generally more complex, but allows lower total capacity. Unlike edge restoration, path restoration can make use of the capacity made available by the working paths that were using the failed edge. This is known variously as *stub release*, freed capacity, bandwidth re-use, or capacity re-use. *Demand splitting* refers to the ability of network equipment to split a demand into smaller, switchable units of capacity (in this investigation, the switching granularity is OC-1). This investigation studies modular capacity planning in a mesh network with demand splitting and stub release using path restoration of all demands affected by a single edge failure.

Due to the complexity of the optimization problem, it is typically decomposed into two stages: first, the working capacity is minimized; then, based on these working paths and modularity restrictions, the spare capacity is minimized. If existing capacity constraints can be ignored in the first stage, then shortest path routing minimizes working capacity. The spare capacity needed for restoration is minimized by determining routes that fill in the gaps left between the working capacity and the modular implementation, plus any additional restoration capacity. The restoration routes can also take advantage of demand splitting in order to minimize the total modular capacity. In this paper, this is termed the *Sequential Capacity Allocation* (SCA) approach.

The *Joint Capacity Allocation* (JCA) approach jointly considers the effects of working and restoration routing. By rearranging the working paths, the benefits of restoration capacity sharing are increased. For example, in Figure 1a, there are two demands, a demand from node 1 to 5 for two units of bandwidth, and from node 1 to 7 for three units. Assume the available modularities are OC-3 and OC-12 and only one fiber-pair per edge is allowed, i.e. an edge will consist of either a single OC-3 or a single OC-12. By routing working paths according to the fewest hops, the minimum total capacity needed is 51 units. However, by taking different working paths, so that restoration capacity is shared, as illustrated in Figure 1b, the total capacity is reduced to 30 units. This simple example illustrates that JCA should achieve solutions that are better than SCA. However, SCA will generally be less complex and thus faster than JCA. Also, it should be noted that working and/or restoration paths may be predetermined. The model used will accommodate any level of predetermined paths. The model will not change any working path that is not directly affected by a failure, thus providing a more easily managed network.

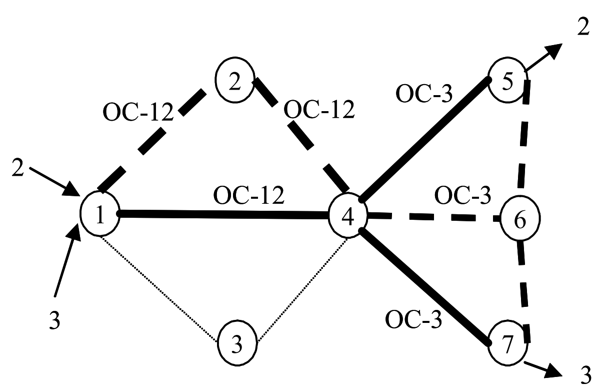

Figure 1a. SCA routing requires 51 total units of capacity.

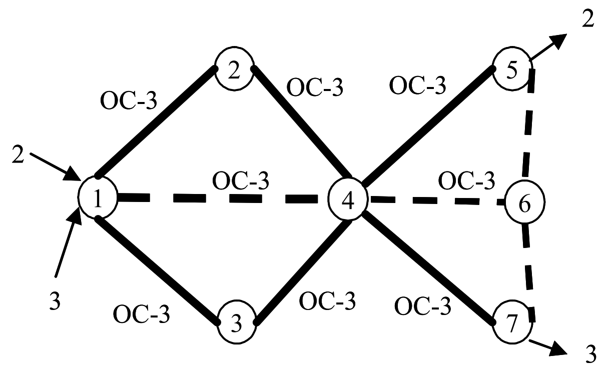

Figure 1b. JCA routing requires 30 total units of capacity.

Legend

Working Path Restoration Path

Figure 1. Effects of Shortest Path Routing versus Joint-Distributed Routing on Total Network Capacity

A thorough investigation of the mesh JCA problem is found in Iraschko et al [12] where results of joint versus sequential mesh network optimization are presented. They found the greatest benefits for joint optimization in span restoration, that is, the per cent capacity improvement of span restorable networks was greater than path restorable networks. In [13] the redundancy of a path-restorable network is found to improve if the demand routing disperses the required capacity throughout the spans of the network, implying increased network connectivity improves total capacity. An essential difference between [12] and this investigation is our use of DOE techniques to analyze modifications to a generic network, rather than comparison test algorithms against a broad mix of network sizes and types. Thus, we investigate the effects of specific network characteristics on capacity (versus comparing six different restoration approaches on five different networks). Also, the effects of modular capacity constraints [7, 16] are included in this investigation.

Murakami and Kim [19] also compared SCA and JCA using an arc-path based LP model. They tested their column-generation solution method using ten randomly generated demand patterns plus a uniform demand as well as two patterns placing larger demands between adjacent nodes. They tested these various demand patterns on two networks: an 11 node, 46 edge network and a 28 node, 90 edge network. On the small network with higher node degree, they found average cost-of-capacity improvements of 7.2% (JCA over SCA). On the larger, sparse network they found average cost improvements of 10%. This is consistent with this investigation's findings that increasing connectivity decreases JCA benefits. Key differences are that [19] did not include modularity effects and the path flow variables were allowed to be fractional, using a round-up post-processing heuristic to achieve integer solutions.

Other approaches to the general problem of capacity allocation in robust networks may be found in [5, 7, 8, 14, 21]. The effects of including modularity in the formulation for finding minimum total capacity in a span-restorable mesh network is investigated in [7].

They reported significant cost savings by using modular joint capacity placement and also noticed that in half the test networks, multiple spans required no capacity and could be eliminated. The effects of graph connectivity on the type of mesh protection and restoration scheme is presented in [8]. They used an extensive testing/validation approach on six different mesh-restoration schemes: 1+1 path protection; span restoration/spare capacity assignment; span restoration joint capacity assignment; meta-mesh; shared backup path protection; and true path restoration. One result is that decreasing the average node degree from 3.2 to 2.2 resulted in an approximate 50% increase in capacity costs. Kennington et al, [14], jointly consider the effects of working and restoration routing in their investigation of the wavelength routing and assignment problem in a survivable wavelength division multiplexed mesh network. They used a successive shortest path approach to generate cycles from link-disjoint paths between demand pairs.

DOE methodology has been suggested [2] as a rigorous approach to testing the interaction effects of a heuristic's parameters. A complete groundwork for the statistical design of experiments, including terminology, can be found in [18]. DOE begins with selection of the significant parameters or factors thought to affect the metrics of interest, e.g. algorithm effect on total modular capacity. This method of testing provides insight into the problem and into further testing directions via the levels and degree of interactions between the parameters. Multi-variate (not simple linear) regressions are easily calculated from the results of a DOE. If the regression accuracy is validated, then quick estimates of the metric of interest at intermediate parameter values are available. To illustrate these concepts and their application to network management, this paper will utilize a complete 2^k factorial experimental design methodology [6]. This approach sets the selected parameters at one of two "extreme" levels with all combinations of settings creating the sample space.

II. Mathematical Description

A. Background

There are two basic approaches used to model the flows in a network, arc-path and node-arc [16]. The arc-path model allows the user control over allowable paths as well as control over the size of the problem formulation. A node-arc model is based on conservation of flow at a node, equating the amount of flow into a node to the amount of flow out of a node. The node-arc model implicitly includes all possible flow combinations, hence, all possible paths are considered, generally allowing lower capacity solutions than arc-path models. However, this comes at the cost of control over the paths in the solution and problem instance size. Therefore, in this paper, as in most others found in the literature, we will use the arc-path model.

B. Notation and Mathematical Model

Models for joint capacity allocation have been presented before [1, 12, 19], however the model used to generate problem instances is presented here for completeness. Additions to the JCA model include capacity modularity, not allowing working path reassignment of unaffected working paths during a restoration, and an objective function that also weights the spare and working capacity. The following notation is defined to present the *joint working-spare capacity allocation problem with modularity restrictions and bandwidth re-use.* Let $[N,E]$ be a network with node set $N = \{1, \ldots, \overline{n}\}$ and edge set $E = \{e_1, \ldots, e_{\overline{m}}\}$ of ordered pairs of nodes. That is, $e_m = (i, j)$ with $i \neq j$ and i,

$j \in N$. A *path* in $[N,E]$ is defined as a sequence $\{i_1, e_{m_1}, i_2, e_{m_2}, \ldots, i_p, e_{m_p}, i_{p+1}\}$ where edge $e_{m_q} \in \{(i_q, i_{q+1}), (i_{q+1}, i_q)\}$ and each edge and each node are distinct. For a path $P^k = \{i_1, e_{m_1}, \ldots, i_p, e_{m_p}, i_{p+1}\}$ we say that the *origin* for P^k is node i_1 and the *destination* for P^k is node i_{p+1}. Let $P = \{P^1, \ldots, P^{\bar{k}}\}$ denote the set of all paths in $[N,E]$. Let o_k, and d_k denote the origin and destination for path P^k.

The working paths comprise the index set $V = \{k: P^k \text{ is a working path}\}$. Let ω_k be the working flow on path P^k. If edge e fails, then service will be interrupted on all working paths that contain e. For each such affected path P^k, we seek to reroute the f_k circuits from o_k to d_k using $[N, E\backslash\{e\}]$. The possible restoration paths comprise the index set $K = \{k: P^k \text{ is a restoration path}\}$. Note that the sets K and V can be equivalent, that is, a single set of enumerated paths can be acceptable for use as either working or restoration paths. Alternatively, if only one working path is specified for each demand pair, then the problem becomes SCA.

Let g_k^e be the flow on path P^k used for restoration when edge e has failed. Let y_m denote the spare capacity allocated to edge $e_m \in E$ with corresponding vector $\mathbf{y}$. Let ω_m denote the working capacity allocated to edge $e_m \in E$ with corresponding vector $\mathbf{w}$. Let ψ_m denote the total modular capacity allocated to edge $e_m \in E$ with corresponding vector $\mathbf{\Psi}$.

The various sets needed for a compact formulation of this problem are defined below. Define the index set of all paths between an origin i and destination j as $O_{ij} = \{k : o_k = i, d_k = j, P^k \in P\}$. Define the set of all node-pairs with demand as $D = \{(i, j) : d_{ij} > 0\}$. Let F^e denote the index set of paths which contain edge e, i.e., $F^e = \{k : e \in P^k, P^k \in P\}$.

Let G^e denote the index set of paths which do not contain edge e, i.e., $G^e = \{k : e \notin P^k, P^k \in P\}$. Let H^{ie} denote the index set of paths which contain edge i, but which do not contain edge e, i.e., $H^{ie} = \{k : i \in P^k, e \notin P^k\} = \{k : k \in F^i \cap G^e\}$. Let I^{ie} denote the index set of paths which contain both edge i and edge e, i.e., $I^{ie} = \{k : i \in P^k, e \in P^k\} = \{k : k \in F^i \cap F^e\}$.

An edge is a fiber-pair that is terminated in a circuit card contained in a cross-connect housed at a node. The number and type of circuit cards determine the total capacity of that edge, i.e. capacity is only available in certain modular amounts. Suppose capacity can only be implemented in fixed sizes from the set $\Pi^e = \{0 = M_0, M_1, \ldots, M_{n_e}\}$. Currently, $\Pi^e = \{0, 3, 12, 48, 192\}$. Let Ω_e^k denote the maximum number of edges of size M_k that can be implemented on edge e, that is, cross-connects have an upper limit on the amount of capacity they can switch and may only terminate certain modularities. Let Φ_e denote the minimum of the maximum number of edges allowed at either end of edge e. Thus, Φ_e is the upper limit on the total number of edges that can be used to implement capacity for an edge e. With these definitions, the set of possible modular capacities for an edge e is: $\Delta^e = \{M: M = \sum_k (M_k s_k), M_k \in \Pi^e, s_k \in \{0, 1, \ldots, \Omega_e^k\}, \sum_k s_k \leq \Phi_e\}$.

Equations (1) – (9) describe the JCA optimization problem as a mixed-integer linear program. The objective is to minimize the costs of the modular bandwidth installed on an edge. (In this investigation capacity economy of scale was not considered, that is, **c** = **1**). Equation (2) states that the total modular bandwidth of an edge *i* is an upper bound on the sum of the upper bounds on the spare and working capacities for edge *i*. Equation (3) requires that the sum of all working paths that use an edge *i* be less that some upper bound for that edge. Equation (4) states that the sum of all restoration paths that use an edge *i* are less than an upper bound plus the amount of freed capacity. Equation (5) requires that all demand be met by choosing flows for working paths. Equation (6) ties the working and restoration together, requiring that restoration flows be generated for the chosen working paths. Equation (7) requires that the total bandwidth on an edge be modular and equations (8) and (9) relax integrality for the non-negative flow variables g_k^e and w_k.

$$\min \mathbf{c}' \Psi \tag{1}$$

s.t.

$$\psi_e \geq \omega_e + y_e, \qquad \forall e \in E \tag{2}$$

$$\sum_{k \in F^e} w_k \leq \omega_e, \qquad \forall e \in E \tag{3}$$

$$\sum_{k \in H^{ie}} g_k^e - \sum_{k \in I^{ie}} w_k \leq y_i, \ \forall i \in E, \forall e \in E \tag{4}$$

$$\sum_{k \in O_{od}} w_k = d_{od}, \qquad \forall (o,d) \in D \tag{5}$$

$$\sum_{k \in O_{od} \setminus F^e} g_k^e = \sum_{k \in O_{od} \cap F^e} w_k, \qquad \forall e \in E, \forall (o,d) \in D \tag{6}$$

$$\psi_e \in \Delta^e, \qquad \forall e \in E \tag{7}$$

$$g_k^e \geq 0 \text{ and integer}, \qquad \forall e \in E, \forall k \in K \tag{8}$$

$$w_k \geq 0 \text{ and integer}, \qquad \forall k \in V \tag{9}$$

Note that because neither the working nor spare capacity is weighted in the objective function (1), they are not necessarily minimized, i.e. only the total capacity is minimized. If total, working and spare capacity were weighted in the objective function then an appropriate balance of working, spare, and total capacity may be achieved, see [9] for a related article on bi-criteria objective functions. The objective function would then take the form:

$$\min \mathbf{c}' \Psi + \alpha'\mathbf{y} + (\mathbf{1} - \alpha)' \mathbf{w} \quad (10)$$

where α is a constant vector between **0** and **1**. If individual edges require different emphases, then α does not have to be a constant vector. To emphasize total capacity over working and spare, the elements of the vector **c** would be given a much larger value than α. To emphasize the minimization of working capacity over spare capacity, elements of α would be given a value less than 0.5.

III. Test Setup and Results

A. Path Enumeration

For comparison purposes, both the SCA and JCA models use the same enumerated path set consisting of the k-shortest paths [13, 17] and the n-shortest paths. The shortest paths are found from Dijkstra's algorithm using hop counts instead of distance. The k-shortest paths are found by iteratively removing all edges that a previous shortest path had found, then running Dijkstra's algorithm on the remaining edges of the network. Note that the disjoint set of paths found in this way is one of several possible disjoint sets and that the paths in the set are not necessarily short. The n-shortest paths are calculated by removing a single edge from the shortest path, running Dijkstra, then replacing the edge. This is done for each edge in the shortest path. Thus, for any edge failure of a shortest path (which is the working path in SCA), there is a mix of short paths and disjoint paths to choose between. Dunn, et al [10] has found that the number of paths in the solution set is a very small subset of the set enumerated for the problem instance and suggests that in practice, the enumerated set can be reduced to a manageable size without dramatically affecting the solution.

B. Experimental Design

Design of experiments is most often associated with quality design and production control. Taguchi methods [22] stress controlled experimentation as part of a robust design. In a 2^k DOE, parameters of interest are identified and two levels for each parameter are determined. The selection of parameters for this experiment are now discussed. Previous work [8] indicates that average network degree and total capacity are inversely related. The amount and number of demands and their distribution obviously has an effect on capacity. An even distribution of equal demands (OC-1) between all nodes is often used in testing [12, 13, 19] . Demand distribution corresponding to hubbing [13] places smaller demands at small nodes, and larger demands at large nodes. The number and type of modular card slots available determines the edge capacity and will therefore have an impact on total capacity. The types of modular cards used here are OC-12, OC-48 or OC-192. The number of cards available per edge is the number of slots. The effect of SCA and JCA on capacity is the fourth parameter. DOE testing will quantify the average main and interaction effects between these parameters on network capacity.

The four parameters are set at either of two levels, creating the 16 experiments of the 2^k factorial DOE. The values for the low and high parameter levels are shown in Table 1. The parameter setting levels for the sixteen tests are shown in Table 2. Note that the test ID corresponds to the settings of the parameters, e.g. test number 101 indicates, respectively, a high average node degree network, with 150 hubbed demands, and a high number of slots per edge.

Table 1. Values for Design of Experiments Parameter Levels

Parameter Level Setting	Average Node Degree	# Demands	# Slots	Algorithm Used
Low (0)	2.56	150 / "Hubbed"	1	Sequential (SCA)
High (1)	3.76	300 / "Even"	3	Joint (JCA)

Table 2. Clarification of Parameter Settings for Each Test Run of SCA and JCA

Test ID	Average Node Degree	# Demands	# Slots
000	Low	150	1
001	Low	150	3
010	Low	300	1
011	Low	300	3
100	High	150	1
101	High	150	3
110	High	300	1
111	High	300	3

The two test topologies (the nodes and edges) are derived from the nationwide Qwest IP and broadband networks (see www.qwest.com and Figures 2 and 3). The nodes were slightly modified so as to represent the 25 U.S. cities with the highest populations, as found in the 2000 U.S. census. The first demand set models hubbing by placing demands between the highest populated cities and all others, e.g. a demand is set between New York City and the remaining 24 cities, then between Los Angeles and the remaining 23 cities, and so on until 150 demands are placed. Demands placed with any of the five largest cities have 2 units of demand added to the origin-destination demand, while the remaining add 1 unit. Thus, the demand amount depends on the size of the end-node cities, ranging from 4 units between two large cities to 1 unit between two small cities. The input data files used are available at www.faculty.bus.olemiss.edu/mlewis/research/data. The second demand set places a demand between all cities (300 total) with the amount held constant at one unit.

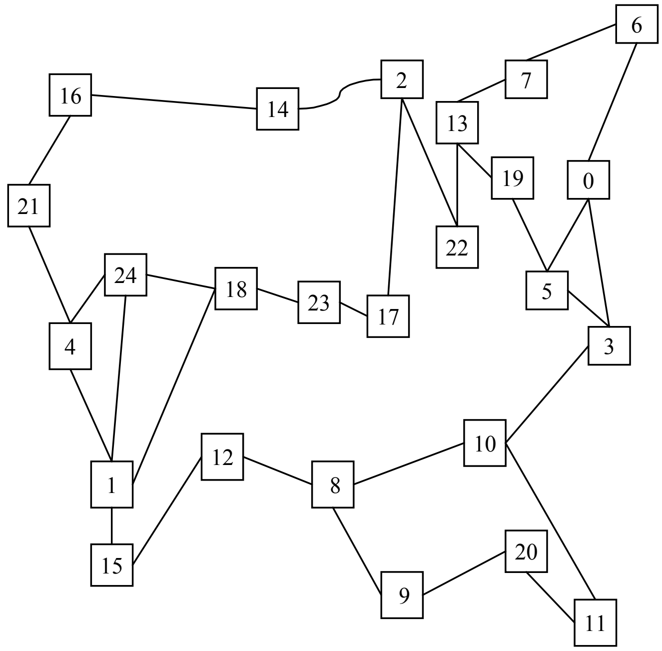

Figure 2. Network with Low Average Degree of Connectivity (2.6)

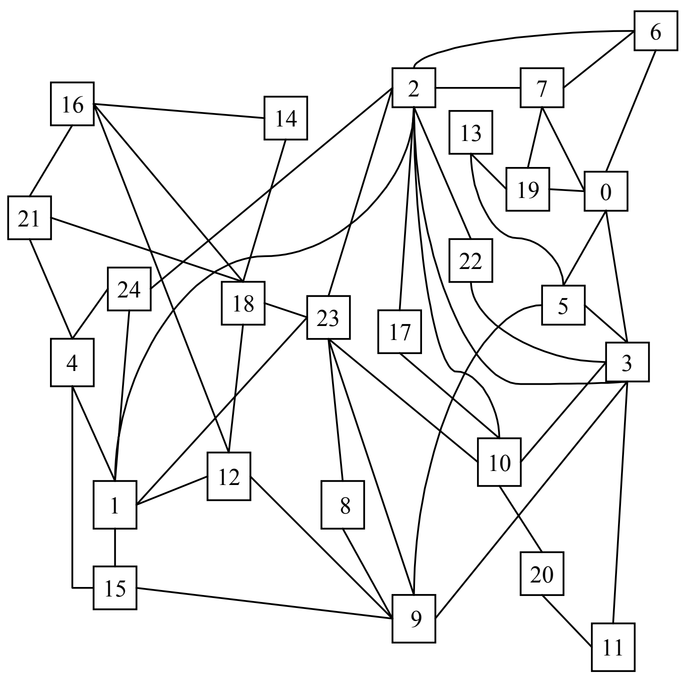

Figure 3. Network with High Average Degree of Connectivity (3.8)

B. Results of Testing

The problems were solved using default settings for branch-and-bound with CPLEX 6.5 run on a Sun Enterprise 3500 with two 400 MHz UltraSparc processors and 1 GB of memory. Parallel processing was not utilized. All problems were either solved to optimality or terminated by CPLEX due to memory limitations (see Table 3).

The test times and optimality gaps are shown in Table 3. The long run times for JCA (2 to 6 days) are due to the difficulty in proving optimality, i.e. the optimal solution is found in a comparatively short period of time, proving it is optimal takes much longer. JCA run times can be improved by reducing the number of path variables and/or removing modularity constraints. The large optimality gaps for the JCA problems having only one fiber-pair per edge indicate that the LP relaxations at the nodes of the branch-and-bound tree do not provide tight bounds. For example, enforcing the requirement to install a single OC-192 when only 60 units are needed creates a large optimality gap.

Table 3. Time and Optimality Gaps

TestID	Time (cpu seconds)		Optimality Gap (%)	
	SCA	JCA	SCA	JCA
000	1	31	0	0
001	2	551	0	0
010	1	337	0	0
011	1	2547	0	0
100	5	184229	0	25
101	103	184412	0	2
110	1	535186	0	28
111	118	535566	0	7

The results of the experimental design are arranged in Table 4 to emphasize the comparison of SCA to JCA. Table 4 illustrates that JCA always improved on total capacity allocation over SCA. This reduction was consistent across all changes in network characteristics, i.e. neither node degree, number of slots, nor demand level dramatically affected the *per cent improvement* of JCA over SCA (see Table 5).

Table 4. Design of Experiments Results

TestID	SCA			JCA			□	□	□
	Spare	Work	Total	Spare	Work	Total	Spare	Work	Total
000	2444	1098	4092	557	1553	3228	1887	-455	864
001	1309	1098	2676	488	1508	2088	821	-410	588
010	2526	1144	4704	600	1600	3372	1926	-456	1332
011	1297	1144	2808	404	1576	2016	893	-432	792
100	853	657	2184	156	760	1308	697	-103	876
101	506	657	1320	131	776	948	375	-119	372
110	685	797	2472	229	980	1860	456	-183	612
111	578	797	1488	166	972	1188	412	-175	300
Averages	1275	924	2718	341	1216	2001	933	-292	717

However, the *total capacity improvement* of JCA over SCA *is* affected by certain network characteristics. Table 6 shows the test results summarized by network characteristic with capacity averaged over the two demand levels. The table illustrates that per cent capacity improvement has small variance, while total capacity improvement steadily decreases. Figure 4 graphs the total capacities of Table 5 to illustrate that as network connectivity and edge slot modularity increase, the total capacity is reduced and the total capacity benefit of JCA over SCA is subsequently reduced. We can conjecture that if a network were totally connected with enough capacity to cover any demand on any edge and the modular capacity installations were as small as the smallest unit of demand, then SCA and JCA would give equivalent capacities. Figure 5 gives an indication of this convergence of SCA and JCA by graphing the differences in spare, working, and total capacities.

Table 5. Total Capacity Improvements

Network Characteristics		Total Capacity averaged over demand levels		□ Capacity	□□Capacity Improvement over SCA
# Slots	Avg Degree	SCA	JCA		
low	low	4398	3300	1098	25%
high	low	2742	2052	690	25%
low	high	2328	1584	744	32%
high	high	1404	1068	336	24%

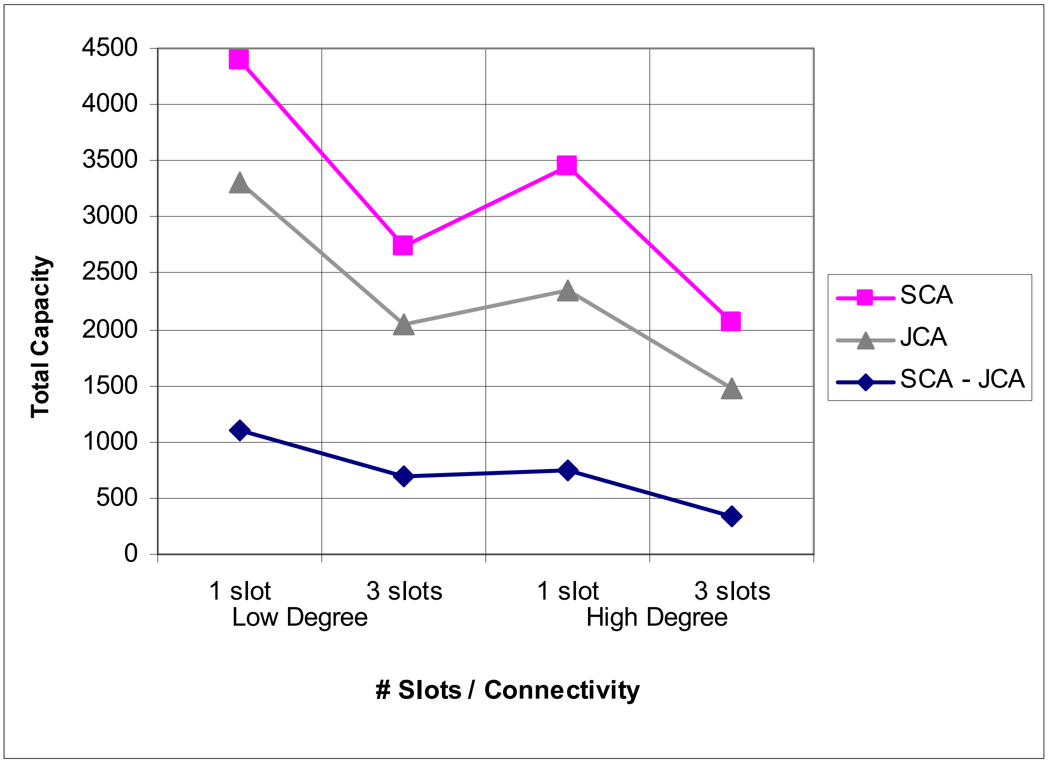

Figure 4. As connectivity increases, total modular capacity requirements decrease and the benefit of JCA over SCA decreases.

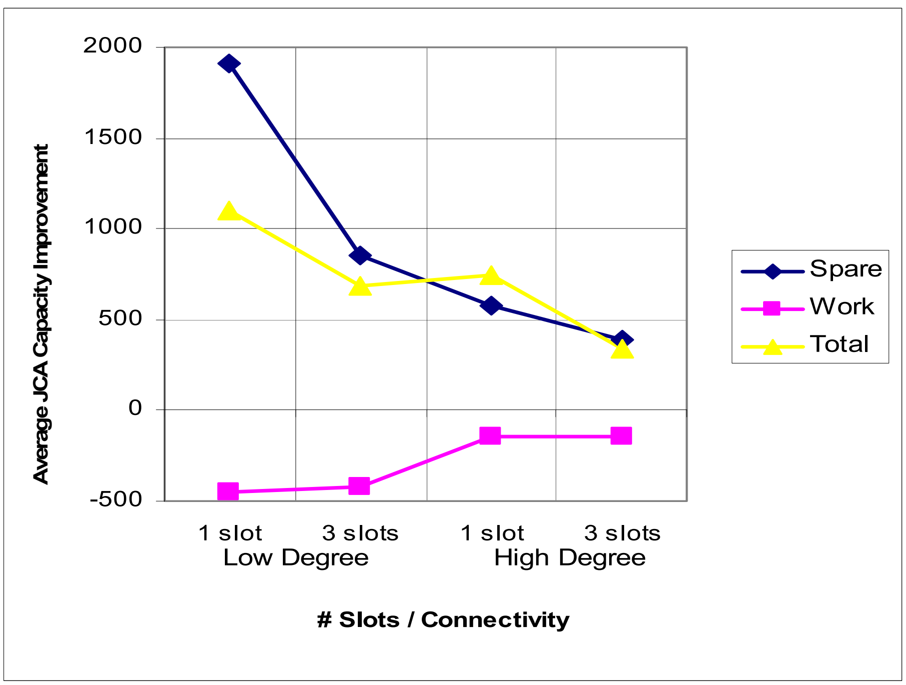

Figure 5. As connectivity and edge modularity/capacity increases, spare, working and total capacity differences converge.

C. Regression Equations

Let the parameters representing number of slots, number of demands, average node degree, and algorithm type be represented as variables x_1, x_2, x_3, x_4, respectively. In calculating the regression, the x variables are set at either -1 to indicate low settings, or at +1 to indicate a high setting. The coefficients for the variables of the multiple linear regression equation derived from the DOE are shown in Table 6. The magnitude of the coefficients for the four parameters indicates their average impact on capacity. For example, the effect of JCA versus SCA on the overall average capacity is 359 units. This is described in the regression as the x_4 coefficient -359, indicating that when x_4 is changed from -1 to +1, then a decrease in capacity of 2*-359 = -718 will occur. Alternatively, the average capacity over all tests is 2360 and when SCA is used, the average is increased by 359 (or decreased by 359 when JCA is used).

The x_3 coefficient of -764 indicates the largest effect on capacity comes from changes in the network degree. Increasing the average node degree from 2.56 to 3.76 reduced capacity by 1528 units or ~50%! Thus, any increased cost associated with additional edges is mitigated by the reduced capacity required. The x_1 coefficient of -543 indicates that changes to the edge capacity/modularity also has a significant effect on capacity. Increasing the mix of available edge capacities helps a network manager tailor capacity to demand levels. Adding card slots to a cross-connect or upgrading a node to handle more and different levels of capacity could be a cost-effective way to reduce total network capacity.

Interaction effects numbers are harder to interpret, for example, the largest interaction is between slots and degree with a coefficient of 183. This indicates that, on average, changing the slots and node degree parameters at the same time and in the same direction (low slots/degree & high slots/degree for both SCA and JCA) increased capacity by 2*183=366 over changing them at the same time in different directions (low slots/high degree & high slots/ low degree for both SCA and JCA). If the two parameters operated independently, then there should be no difference in the two averages, because both are averages of either of the two parameters at low and high settings, over both algorithms. Experimental design results are often used as a guide for further investigations. For example, more testing to clarify the interaction between edge modularity/capacity and average node degree, or between algorithm type and node degree. However, that extends beyond the scope of this investigation, which will next concentrate on validating the regression equation.

Table 6. Regression Equation Coefficients For Estimating Capacity Based On Variable Network Parameters

Variable Name	Terms	Coefficient
Slots	x1	-543
Demand	x2	129
Degree	x3	-764
Algorithm	x4	-359
Interactions	x1 x2	-71
	x1 x3	183
	x2 x3	-27
	x1 x4	102
	x2 x4	-21
	x3 x4	89
	x1 x2 x3	17
	x1 x2 x4	5
	x1 x3 x4	0
	x2 x3 x4	63
	x1 x2 x3 x4	-29
Average Effect		2360

The regression equation above can be checked by comparing the estimated effect on capacity of intermediate parameter settings to the optimal levels. The middle level of demand is based on increasing the number of demands to 225 (midway between 150 and 300), and decreasing the hubbing so that the three (versus 5) largest cities get larger demand.

Table 7 shows the results of these tests and the per cent error between estimated and optimal capacity levels. The table indicates that the regression from the DOE estimates the capacity at intermediate level of demand (at varying degree and slots) to within ~5%. An alternative regression approach that would be used if DOE methods were not utilized is to average the capacities at low levels of demand, then at high levels, then interpolate for total capacity at intermediate levels of demand. That approach averages out the effects of the other network parameters and generated errors in estimation ranging from -93% to 36% for SCA (versus -9% to 0%) and -38% to 53% for JCA (versus -1% to 10%).

Although the errors for the other two parameters are larger than those for demand, the estimates are consistently higher, e.g. the SCA capacity estimate at intermediate node degree was consistently ~20% too high. Thus, the regression recreates the basic shape of the relationship, but needs to be shifted down (see Figure 6). Simply reducing all SCA intermediate degree capacity estimates by 20% will shift the curve to fit the results nicely. A simple shift to acquire a good fit indicates that while the relationship estimate is accurate, the average capacity (2360 over all the tests) is estimate is not. The error in estimating the average indicates that the node degree / total capacity relationship is probably non-linear. The errors in estimation for JCA at intermediate node degree indicate that both the average and relationship estimates are off.

Levels of demand are typically stochastic variables. Network changes to edge capacity and/or topology occur with acquisitions, restructuring, technology changes, etc. A regression equation to quickly obtain capacity estimates should be useful to managers of large telecom transport networks. Possible uses include quick answers to "what-if" scenarios in purchasing/leasing capacity to match changes in demand. Spreadsheet software such as @Risk and Crystal Ball can make use of fast and accurate estimates during simulation modeling and stochastic decision making. Salesmen who are asked questions about the effects of demand changes on the equipment used in a customer's network could provide quick estimates of equipment needed based on capacity changes estimated. Instead of total capacity, similar multi-variate linear regressions can be calculated and verified for individual edge capacities.

Table 7. % Errors When Estimating Total Capacity Based on an Intermediate Parameter Settings

Network Characteristics			SCA estimate	SCA actual	SCA % Error	JCA estimate	JCA actual	JCA % Error
Demand at 225	Low Degree	Low slots	4398	4668	-6%	3300	3336	-1%
		High slots	2742	2868	-5%	2052	2076	-1%
	High degree	Low slots	2328	2328	0%	1584	1428	10%
		High slots	1404	1536	-9%	1068	1128	-6%
2 slots	Low Degree	Low demand	3384	3096	9%	2658	2364	11%
		High demand	3756	3228	14%	2694	2364	12%
	High degree	Low demand	1752	1368	22%	1128	1020	10%
		High demand	1980	1572	21%	1524	1212	20%
Degree at 3.32	Low Demand	Low slots	3138	2640	16%	2268	1332	41%
		High slots	1998	1560	22%	1518	1128	26%
	High Demand	Low slots	3588	2856	20%	2616	1812	31%
		High slots	2148	1668	22%	1602	1356	15%

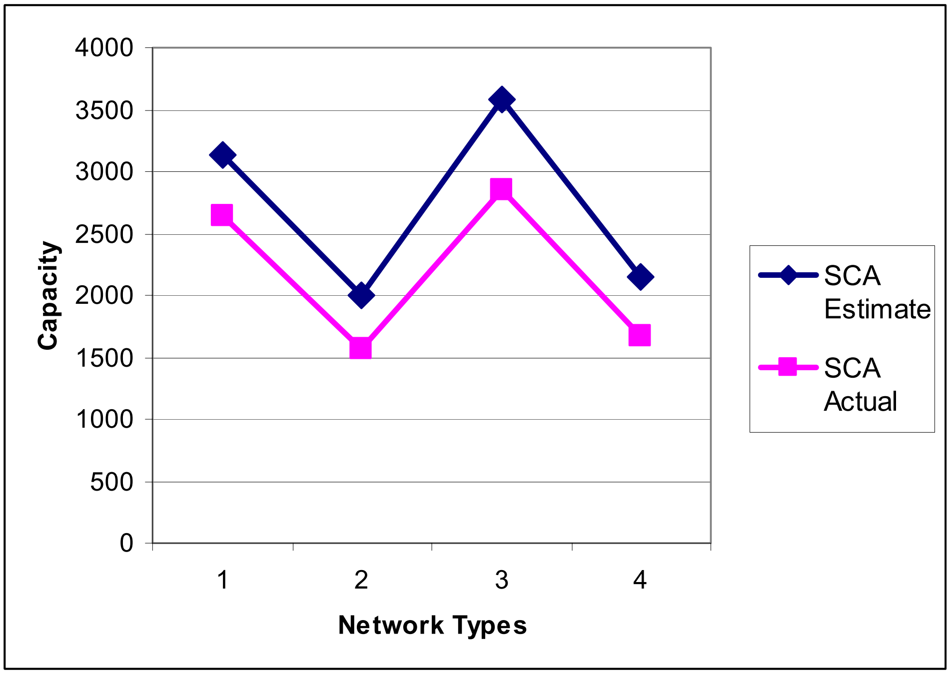

Figure 6. SCA Capacity Estimates At Intermediate Network Node Degree

IV. Conclusions

In this investigation, a design of experiments methodology was used to control and quantify the main and interaction effects of certain network parameters on total network capacity. Using a single generic network, modified eight different ways, it was found that JCA had an average 26% improvement over SCA, with changes to network connectivity and demand configurations having little effect on the per cent improvement in capacity. However, as network connectivity and edge capacity/modularity increases, the total capacity benefits of JCA decreased. It is conjectured that SCA and JCA capacity allocation will converge at the extreme of connectivity and edge capacity/modularity.

The experimental design method also yielded multi-variate regressions whose coefficients provided insight into the interactions between network parameters. As noted above, the regression indicated that network degree and edge capacity/modularity do not operate independently in affecting total capacity. The regression equation was tested for validity and was found to be an accurate indicator (average 5% error) for estimating capacity at an intermediate demand. The regression accurately estimated the relationship of average node degree to total capacity for SCA, but consistently overestimated the average capacity, resulting in consistent errors in estimation and indicating a non-linear relationship.

References:

M. Abou-Sayed, J. Kennington, and S. Nair, "Joint working and spare capacity assignment in a link restorable mesh network," Department of Computer Science and Engineering, Southern Methodist University, Dallas, TX, Technical Report 96-CSE-16, 1997.

R. Barr, B. Golden, J. Kelly, M. Resende, and W. Stewart, "Designing and reporting on computational experiments with heuristic methods," *Journal of Heuristics,* vol. **1** , pp. 9-32, 1995.

S. Cosares and I. Saniee, "An Optimization Problem Related to Balancing Loads on SONET Rings," *Telecommunication Systems*, vol. **3,** pp. 165-181, 1994.

S. Cosares, D. Deutsch, I. Saniee, and O. Wasem, "SONET toolkit: A decision support system for designing robust and cost-effective fiber-optic networks," *Interfaces,* vol. **25** , pp. 20-40, 1995.

T. Craniac and M. Gendreau, "A simplex-based tabu search method for capacitated network design," *INFORMS Journal on Computing*, vol. **12,** pp. 223-236, 2000.

R. DeVor, T. Chang, and J. Sutherland, *Statistical Quality Design and Control: Contemporary Concepts and Methods*. New York, NY: MacMillan Publishing Company, 1992.

J. Doucette, W. Grover, and R. Martens, "Modularity and economy-of-scale effects in the optimal design of mesh-restorable networks," *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 226-231, 1999.

J. Doucette and W. Grover, "Comparison of mesh protection and restoration schemes and the dependency on graph connectivity," presented at the 3rd International Workshop on the Design of Reliable Communication Networks (DRCN), Budapest, Hungary, 2001.

J. Doucette, W. Grover and T. Bach, "Bi-criteria studies of mesh network restoration path-length versus capacity tradeoffs," presented at OSA Optical Fiber Communications Conference, Anaheim, California, March 2001.

D. Dunn, W. Grover, and M. Macgregor, "Comparison of k-shortest paths and maximum flow routing for network facility restoration," *IEEE Journal on Selected Areas in Communications*, vol. **12**, pp.88-99, 1994.

O. Goldschmidt, A. Langier, and E. Olinick, "SONET/SDH ring assignment with capacity constraints," to appear in *Discrete Applied Mathematics*, circa 2002.

R. Iraschko, M. MacGregor, and W. Grover, "Optimal capacity placement for path restoration in STM or ATM mesh survivable networks," *IEEE/ACM Transactions on Networking,* vol. **6**, pp. 325-336, 1998.

R. Iraschko and W. Grover, "A highly efficient path-restoration protocol for management of optical network transport integrity," *IEEE Journal on Selected Areas in Communications*, vol. **18,** pp. 779-794, 2000.

J. Kennington, E. Olinick, A. Ortynski, and G. Spiride, "Wavelength routing and assignment in a survivable WDM mesh network," Southern Methodist University, Dept. of Engineering Management, Information and Systems, Dallas, TX, Technical Report No. 00-CSE-6, 2001, to appear in Operations Research.

J. Kennington, "A survey of multicommodity network flows," *Operations Research* , vol. **26,** pp. 209 – 236, 1978.

J. Kennington and M. Lewis, "The path restoration version of the spare capacity allocation problem with modularity restrictions: models, algorithms, and an empirical analysis," *INFORMS Journal on Computing,* vol. **13,** pp. 181-190, 2001.

M. MacGregor and W. Grover, "Optimized k-shortest paths algorithm for facility restoration," *Software - Practice and Experience* , vol. **24,** pp. 823-834 , 1994.

R. Mason, R. Gunst, and J. Hess, *Statistical Design and Analysis of Experiments*. New York, NY: John Wiley and Sons, 1989.

K. Murakami and H. Kim, "Joint optimization of capacity and flow assignment for self-healing ATM networks," in *Proceedings of 1995 IEEE International Conference on Communications*, 1995, pp. 216-220.

Y. Myung, H. Kim, and D. Tcha, "Optimal load balancing on SONET bidirectional rings," *Operations Research,* vol. **45,** pp. 148-152, 1997.

C. Shyur, T. Lu, and U Wen, "Applying tabu search to spare capacity planning for network restoration," *Computers and Operations Research* , vol. **26,** pp. 1175-1194, 1999.

G. Taguchi, A. Elsayed, and T. Hsiang, *Quality Engineering in Production Systems*. New York: McGraw-Hill, 1989.

Using the CPLEX callable library. ILOG, Inc. Incline Village, NV, 1997.