Multiple Criteria for Evaluating Survivability in a Telecommunications Network

Steven Cosares

Dept. of Business Computer Information Systems
Hofstra University, Hempstead, NY 11549

Abstract

Survivability, which is a measure of the extent to which the network demands can be satisfied, even under the risk of link or node disruption, is a desirable quality that rivals cost as the primary criterion by which network planning proposals are compared. When *all* of the demands can be satisfied despite *any* single network failure event, the term "100% Survivability" applies. However, when some demands are left unserved during some failures, a well-defined value measuring the level of survivability is elusive. We describe a number of schemes that may be applied in this case. They consider the protection provided to demands in their working routes to reduce network vulnerability and the restoration mechanisms in place to for quickly re-routing network traffic in response to a disruption. A planner can apply one or more of the measurement schemes to identify vulnerable demands and/or locations in a network that require additional capacity. Proposed formulas can be used to compare plans for the level and/or type of survivability they provide. The formulae can also be used to define unambiguous quantifiable survivability constraints on the solutions generated by some mathematical programming-based model for planning. These measures are consistent with schemes that measure other forms of network robustness, like insensitivity to demand randomness.

Key Terms: Network Planning, Survivability, Multi-criteria Decision Making

1. Introduction

A network that supports high capacity traffic, for telecommunications or power distribution, must be designed in a way that makes it robust to the potential damage from unforeseen events like a cut in a link or a breakdown of some network equipment. A network having *survivability* is capable of satisfying the demand for the point-to-point services expected of it, despite the potential for such disruptive events. This is achieved when the network is planned with sufficient extra capacity, is multi-connected, and has the ability to immediately re-route traffic, if necessary, to avoid any failed network locations. In an age where customers are increasingly dependent on the reliability of their networks, (consider for instance the consequences if 911 service in an area were suspended for some length of time or the potential fallout from another power outage like the one that hit northeastern U.S. in 2003), the level of survivability is probably as important a consideration to a network planner as the total cost of the network

components. For telecommunications networks, a planner has a number of options, e.g., using equipment and architectures based on the Synchronous Optical Network (SONET) and Wavelength Division Multiplexing (WDM) technologies, for configuring a network with sufficient capacity and traffic-switching capability to provide effective and cost efficient survivability. (See [1] and [2] for detailed descriptions of the commonly used "self-healing" architectures based on the technologies). In these networks, when a cable or equipment failure is detected, the (pre-designed) plan for restoring service is automatically and quickly implemented.

A typical goal for a network planner is to design a network that is cost-effective while providing "100% survivability", where *all* of the point-to-point demands (within reason) can be satisfied despite *any* potential network failure event. To develop a plan with such guarantees on service reliability would require the pervasive application of the self-healing architectures referred to above and would likely require the assistance of some decision support systems like the ones described in [3] through [6]. In cases where it is appropriate to consider alternative levels of survivability, e.g., if the 100% level is prohibitively expensive, the planner would have to evaluate the tradeoff between cost and vulnerability. The planner could save money, for instance, by not protecting every demand or hedging against every possible network failure scenario. He or she may prefer a scheme that provides high levels of survivability where it is needed most and allows lower levels elsewhere. To be helpful, a decision support system for network planning would correspondingly have to provide additional functionality, as described in [7]. As we shall see, candidate network planning solutions would have to be evaluated, compared, and selected based on multiple criteria regarding their survivability, as well as on their cost.

To compare a pair of survivable plans for their associated capital costs, i.e., the cost of new infrastructure, cabling and equipment, (e.g., for multiplexing, switching, signal regeneration and termination), requires a fairly straightforward process of accounting for the required purchases identified during the network design process. However, it is a much more daunting task to evaluate and compare plans based on the level of survivability they provide. A percent value that evaluates partial survivability in a plan that lets some traffic be lost during some failures is elusive, (except when 100% survivability is provided - the only point-to-point demands that are ever lost are those that terminate at a failed node). This is because of a couple of factors. First, survivability in a SONET or WDM network is achieved through the use of a variety of

techniques, including diverse routing, dedicated protection switching, and the use of self-healing architectures that provide shared capacity for re-routing traffic around failed locations. Thus it is difficult to draw a direct comparison between a pair of plans that employ different mixes of these techniques. Second, the calculation for the level of survivability in a plan could be based either on the vulnerability of one or more of the demands or on the potential severity of one or more of the failure events or on some mix of the two. So a given plan could generate a whole vector of reasonable values representing its level of survivability; the planner would have to identify the most important vector component or mix of components (decision criteria) before being able to directly compare potential network solutions.

The remainder of the paper addresses the question, "What value or values accurately reflect the level of survivability provided in a given network plan?" Though we focus on telecommunications networks, the approaches we develop to address the question may be applied to other types of networks that support the secure delivery of commodities. We provide details about different options available to measure (partial) survivability and describe where they may be applied in the network planning process. For a proposed plan having a clear policy for demand protection and restoration, calculations can be performed for each demand and/or for each network failure scenario. These values can be used to identify which locations in the network are particularly vulnerable or which network failure scenarios are particularly severe. The formulae can also be used to define unambiguous quantitative survivability constraints on the planning solutions generated by mathematical programming model for network planning. Some aggregation of these measurements could be calculated to provide an overall (percentage) level of survivability for the network. A planner would selected the measurement scheme that most consistently models his or her most prominent decision criterion for the network. This single metric may be used to compare plans and to evaluate the tradeoff between survivability and other decision criteria like total cost.

2. Definitions and Notation

In order to develop a clear-cut evaluation of network survivability (expressed as a percentage), it is assumed that the demands on the network are known. For stochastic planning models that embrace demand forecast uncertainty this assumption could be adapted, as will be described

later. Each of the *K* point-to-point demands is represented as a number of units of capacity required, *d(k),* between some pair of locations in the network. A demand is satisfied when its units are assigned to one or more paths having sufficient capacity. The link *(i,j)* in the network corresponds to one or more transmission systems between locations *i* and *j* and/or shared architectures covering, like self-healing rings having terminals at *i* and *j.* The capacity of a link is dependent on the capacities of the systems incident there. If a demand unit is assigned to a path of length two or greater, then some cross-connection equipment must be present at the intermediate node(s) to allow switching from one link to the next. (The term "mesh" is often used to describe an architecture consisting of a highly interconnected set of links that support multiple paths between locations. If they have sufficient capacity, meshes can be very effective in providing survivability.) The paths used by a demand during normal operating conditions are referred to as "working" paths. We assume that the network has at least the capacity necessary to support the assignments to working paths.

Figure 1 illustrates the assignment of a demand of ten units to working paths consisting of interconnected systems, each, as we shall see, having a different ability to provide survivability.
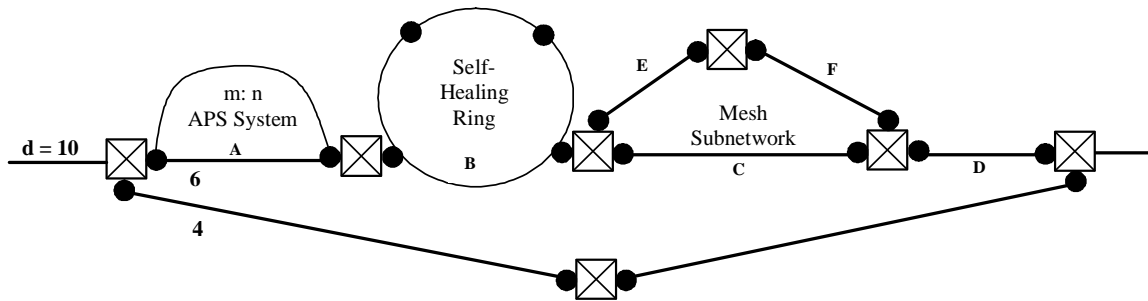

Figure 1

Four units of the demand are assigned to the lower path, which is physically diverse from the upper. The systems comprising the links in the upper path, A, B, C and D, are assumed to have sufficient capacity to accommodate the six units assigned to it.

To account for the survivability provided in a plan, it is necessary to enumerate all of the *F* potential failure scenarios. Each failure scenario *f* corresponds to the loss of some capacity in the network due to some unforeseen disruption. The planner has the option of listing only those

failures he or she considers sufficiently likely and potentially severe. Although each failure may impact multiple links and/or nodes, it is a common approach in survivable planning to assume that each failure corresponds to the disabling of a single link or node in the network. We assume that the network plan, by virtue of its equipment and architecture placements, includes some plan for the survival of the demands for each of the F failure scenarios identified.

There are a variety of ways in which some or all of the units of a demand k would survive a failure. In one scenario, the failure would have no impact on the demand, so no units are lost. In our example, if system represented by link E were disabled, then the ten units of demand would be unaffected. We point out that by splitting the demand over the two diverse paths guarantees that at least four units of the demand will always avoid any single link or node failure. In a second scenario, a link or node affected by the failure may correspond to a system that has "automatic protection switching" (APS). That is, the system has dedicated (diverse) protection capacity that is automatically switched to the moment the failure is detected. In our example, the system represented by link A is equipped with m:n APS that provides m units of dedicated protection capacity for every n units of working capacity. If a 1:2 system were in place then three of the six units of demand would survive the failure of link A. In a third scenario, the failure may occur within a "self-healing" architecture like a SONET ring, which would automatically switch traffic in a direction that avoids the failure. In our example, the demand would survive any failure in the ring, except a breakdown of the node between A and B or the node between B and C. In the final scenario, the link or node failure may occur within a system that is not equipped with pre-defined protection. If the network has sufficient additional capacity and it has the ability to quickly switch traffic to alternative paths when the failure is detected, (e.g., it employs mesh-based architectures), then some units of the demand may still be satisfied. The network plan would contain provisions for how demand units are to be feasibly re-assigned to back-up paths for each failure scenario of interest. In our example, if link C were disabled, it may be possible to switch some demand units directly to links E and F, if capacity were available at both links. Alternatively, it may be necessary to salvage some or all of the demand by routing across the lower path to the extent that the link capacities allow.

Thus, the means of providing survivability can be classified under the terms, "route diversity" (scenario 1), "demand protection" (scenarios 2 and 3), or "demand restoration" (scenario 4). The

first two categories are "preventative survivability" measures, which serve to minimize the vulnerability of the demand in its working path assignments to the potential failures. The third category can be called "reactive survivability" because it is put into action to address demand that would otherwise be lost to the failure.

For a given network failure f and demand k, we define $P(k,f)$ to be the proportion of demand $d(k)$ that survives from the preventative measure. Generally speaking, the values of $P()$ can be kept high through the placement of self healing architectures, systems with dedicated protection, and diversity in the working path assignments. We define the overall survivability measure $S(k,f)$ to be the proportion of demand $d(k)$ that is satisfied due to either preventative or reactive measures after failure f. Note that 100% survivability occurs if $S(k,f) = 100\%$ for all k, f.

For the example in Figure 1, suppose that the transmission system represented by link A provides 1:2 protection, that the lower path has one unit of spare capacity that can be used for protection, and that links E and F have two units of spare capacity. Then if link A were to fail, three units are protected by the system and one unit can be re-routed over the lower path, so $P() = 70\%$ and $S() = 80\%$. For a failure in the ring represented by link B all of the traffic is protected by the ring, so $P() = 100\%$ and $S() = 100\%$. If the system represented by link C were to fail, two units would be restored over links E and F and one unit would be restored using the lower path, so $P() = 40\%$ and $S() = 70\%$. For a failure of link D, only the lower path can be used, so $P() = 40\%$ and $S() = 60\%$. For the failure of link E, $P() = 100\%$ and $S() = 100\%$.

Since some demands are left unserved for some failure scenarios, the plan does not provide 100% survivability. So what percentage does it provide?

3. Survivability Measurements

The remainder of the paper discusses a number of options for organizing the values of $P(k,f)$ and $S(k,f)$ into summary statistics that would help a planner evaluate the network and identify where new equipment and fiber purchases are likely to improve the survivability. We posit that the $2*K*F$ individual values for $P()$ and $S()$ have little utility, except as a diagnostic or control for a specific demand during a specific failure scenario. To develop more meaningful information would require appropriate aggregations of the values.

Low values for *P()* or *S()* across a particular failure scenario, e.g., a link cut, may motivate the planner to place additional systems or dedicated protection capacity at that link or to add capacity to alternative working paths and re-assign some demands to them to ease the burden on that vulnerable link. Low values for the *P()* values associated with a particular demand *k* would indicate that the demand is highly exposed to certain failures. This often occurs when most of the *d(k)* units traverse the same links. If the working paths for the demand are diversified, i.e., the demand units are spread out over multiple paths wherever possible, then a majority of the demand will avoid any single failed link or node in the network. Low values for *S()* over the same demand would indicate that the restoration scheme is doing little to rectify this deficiency in planning. To improve survivability in this case would require the placement of additional systems, perhaps in new locations to increase the potential for route diversity.

Planners employing route diversity often require that a demand be evenly split among paths that are completely disjoint. We point out that reasonable goals for survivability can be met without so stringent a rule. We also point out that for networks having a fixed topology, there are limits to the level of diversity a planner can provide to the working path assignments. Hence, it may not be feasible to place some specific constraints regarding minimum values for some *P()*.

3.1. Demand-based measurements

Our discussion, henceforth, will focus on developing summary statistics based on the *S()* values. As discussed above, useful information could be drawn from performing similar calculations on the *P()* values as well. If a planner deems that certain demands do not require protection, they may be excluded from the evaluation of *S()*. For the remaining demands, it might be helpful to group the values by demand. This would give information about how sensitive each demand is to the prominent failure scenarios identified by the planner. Figure 2 provide a sample graph that may be generated for some demand *k* in a network planning instance. It is obtained by placing the values of *S(k,*)* in decreasing order.
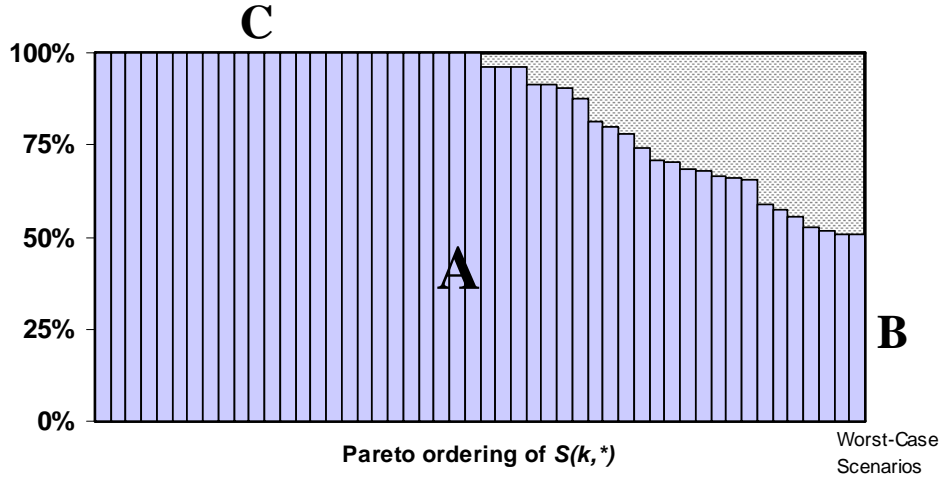
Figure 2. The Protection/Restoration of demand k

The region in the upper right corner represents when units of the demand do not survive a failure. So if demand k were provided with 100% restoration in all F failure scenarios, the region on the upper right would not exist. As drawn, the chart reflects the assumption that any failure is equally likely. If the planner has some probability information detailing the relative likelihood of each failure scenario, the widths of the corresponding rectangles could be adjusted accordingly.

The chart shows that there are at least three alternative ways to measure how well demand k survives network failures. Area A corresponds to the percentage representing the average proportion of $d(k)$ served, over the set of scenarios. Its value is $A(k) = \Sigma_f S(k,f) / F$. This value can be interpreted as the average level of insensitivity of demand k to the failures. The length indicated by B corresponds to the percentage of demand k that is served in the worst-case failure scenario. (Different demands may have a different failure scenario as its worst case). So $B(k) = \min_f S(k,f)$. This value can be interpreted as the minimum guaranteed level of survivability provided to the demand. The length indicated by C represents the proportion of scenarios in which 100% protection is provided to the demand, so $C(k) = | \{ f | S(k,f) = 1.0 \} | / F$. It can be interpreted as the probability that demand k would completely survive, should some random failure occur. We point out that the values of A, B, and C are likely to be very different, except when 100% survivability is provided. Any of these measures is useful in some network-planning context. Depending of the focus of the planner, one would be more prominent and would be used

as a criterion to evaluate the overall plan. This would also dictate the particular methods that would need to be employed to improve demand survivability. For instance, if $A(k)$ is deemed too low, then it may be increased by providing dedicated protection to or improving the diversity of the working paths for the demand. If $B(k)$ is too low, then the planner must focus on the worst case failure scenarios.

3.2. Scenario-based measurements

Additional information would be provided if the data were organized by failure scenario. By observing which demands are not served because of a particular failure, the planner is provided with a view of the failure's relative severity. The following graph is generated for failure f by placing the values of $S(*,f)$ in decreasing order.
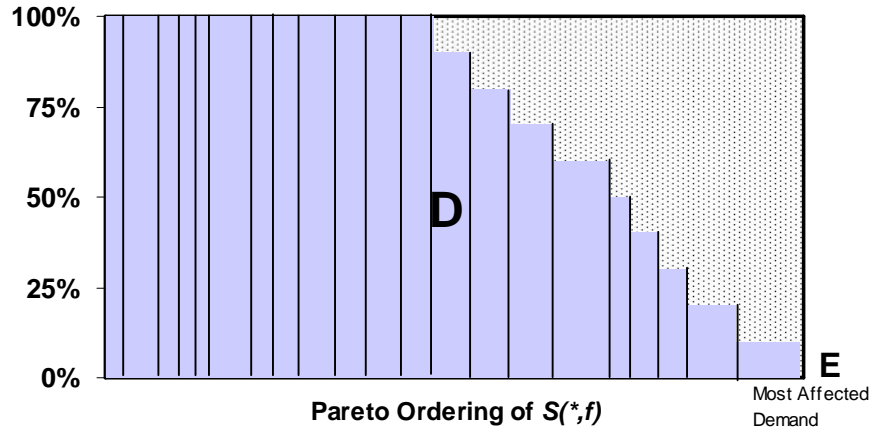

Figure 3. The impact of failure scenario f

In this diagram, the bars have widths that are proportional to the number of units in the demand. The area represented by D is the proportion of the total demand that does not survive failure f. Its value is given by $D(f) = \Sigma_k S(k,f)*d(k) / \Sigma_k d(k)$. It can also be interpreted as the weighted average survivability from failure f, where the weights are based on the demand levels. The length represented by E is the minimum survivability from failure f, given by $E(f)=\min_k S(k,f)$. Any plan with service guarantees would seek to keep this number large. If the region in the upper right were triangular in shape, (as it is in the above illustration), that would indicate that the recovery plan is not equitable across demands. We point out that a major challenge in the design

of optimization algorithms for partial survivability is to establish a mechanism that chooses which demands to restore and which demands to leave unserved after a failure. A low value for either *D(f)* or *E(f)* would indicate to the planner that survivability would be improved if there were a greater number or more capacious paths that can serve as a backup to the affected link or node.

We point out that the graph obtained by ranking the *P()* values associated with failure *f* would be similarly valuable to a network planner since it provides a view of how much demand is impacted by *f*. By placing both graphs on the same axes, one could then measure the effectiveness of the restoration plan.

3.3. Application to mathematical programming models

The survivability measurement schemes described above help the planner evaluate some existing plan for network expansion, demand routing, protection and restoration. The diagnostics help identify which demands are particularly vulnerable and which failures are particularly severe, so that adjustments to the plan can be implemented as needed. In situations where planning decisions have yet to be made, some mathematical modeling approach may be employed. The model would help the planner determine which systems to purchase, their capacities, and where to place them. For ease of notation, let us represent these decisions by the vector *z*. Additional decision variables will determine how to assign demands to routes during normal operating conditions; let us represent these variables by the vector x^N. Suppose the variables are governed by the constraints, $p_k^N(\mathbf{x}^N) = d(k)$ *for all k*, stating that the working routes satisfy the demands. Though these constraints are often represented as mixed integer linear system, we use the functional notation for greater generality. Another constraint system, $A^N(x^N) <= C^N(z)$, states that the routing assignments do not violate the capacities of the transmission and switching systems placed.

During failure scenario *f*, some of the routing decisions are unchanged because they are unaffected or use only protected systems. The assignment of some affected demands to alternative paths for restoration can be represented by the vector y^f. The constraint accounting for this activity is $p_k^f(\mathbf{x}^N) + q_k^f(\mathbf{y}^f) = d(k) - l(k, f)$ *for all k and f,* where *l(k,f)* represents the number

of units of demand *k* that would be left unserved. The first term on the left hand side corresponds to the protection provided demand *k* during scenario *f,* which we previously denoted by $P(k,f)$. The sum of the two terms on the left correspond to $S(k,f)$. Finally, since the failure may decrease the capacity of some systems, thus limiting the amount of restoration that can be provided, we include the constraints set, $A^f(x^N, y^k) <= C^k(z)$.

Thus the model provides a mechanism to control the level of survivability offered in some optimal, (e.g.,, minimum cost), solution. Formulae representing $A(k)..C(k)$ and/or $D(s)..E(s)$ could be used to represent unambiguous mathematical constraints on the exposure of some or all demands or on the level of severity allowed to specific failure scenarios. (See e.g., [8], [9]). For example, the constraint, $E(f) \geq 0.9$ *for all f,* which indicates that the network must be designed so that no more than 10% of *any* demand is lost in any failure scenario, can be represented as $l(k,f) <= (.1)d(k)$ *for all d and k.* The constraint, $D(f) \geq 0.9$ *for all f,* which indicates that no more than 10% of *the total* demand can be lost in any failure can be represented as $\Sigma_k l(k,d) <= (.1)\Sigma_k d(k)$.

3.4. Unified network survivability measurements

To develop a single value that represents the overall network survivability, e.g., that represents some overall objective for the plan, requires further aggregation of the P() and *S()* values. There are a number of alternatives for this as well. The formulae provide additional options for how the planner may constrain the planning solution generated by the mathematical programming model. The formula that best represents the requirements of the planner may be selected as the objective function for the model, replacing or augmenting the objectives associated with cost.

Unlike the term, "100% Network Survivability", which has a single unambiguous meaning, the term "90% Network Survivability" is subject to a wide variety of interpretations. As we have seen, a network that allows for the loss of some demand units generates a large number of different values that represent the protection and restoration provided. Nonetheless, planners often want some broad, single measurement to evaluate network survivability. Such a value would allow the planner to rank plans based on this single value or to evaluate the trade-off between the survivability measure and some other measure, like the total cost of the network.

The most obvious choice for a single network measure would be $S_{min} = \min_k \min_f S(k,f)$. This would be interpreted and the guaranteed level of survivability provided to *every* demand under *any* failure. Constraints based on this measure are most restrictive. Alternatively, the network measurement could be based on one of the aggregate formulas for the demands or scenarios described earlier. In [8], the authors suggest using the value, $D_{min} = \min_f D(f)$. This can be interpreted as the guaranteed minimum survivability provided to the *total* demand for any failure. Setting a minimum constraint on this value, e.g., 90%, allows some individual demands to be greatly impacted by a failure, as long as the total demand throughput is high enough. Similarly, a constraint on B_{avg}, which is the weighted average of the $B(k)$, using the demand levels as weights, allows certain failures to be quite severe as long as the overall guarantee supplied to the demands is high enough.

Even more freedom in the planning solution is allowed if the overall survivability measure is based on some average of the $S()$ values, e.g., the average of $D(f)$, $\Sigma\, D(f) / F$, which is equal to A_{avg}. This value measures the average survivability provided to the total demand. Setting a high target for this value would still allow some individual demands to be completely devastated by some failures, as long as there are a sufficient number of cases where most of the demand is recovered.

4. Conclusions: Planning under Demand Uncertainty

As we have seen, if a planner wants a network that has at least 90% network survivability, he or she may be placing a constraint on every individual demand or placing a single constraint corresponding to some aggregate measure over the demands. Furthermore, the constraint could be associated with the protection provided in some worst-case failure scenario or it can be associated with some average-case failure event. Alternatively, the intent of this percentage may be to constrain the amount of overall damage caused by any single network failure event. Depending on the planner and the network, any of these interpretations is valid. The formulae developed in this paper provide the planner with a means to express any such constraints unambiguously. We point out that the measurements for survivability we seek are assumed to have values that are less than or equal to 100%. We are led to wonder whether it would be

meaningful, e.g., for a demand that is considered extremely important, to constrain some measure to have value that is greater than 100%.

Attempts by the planner to increase $P()$ would likely be more expensive than increasing $S()$ because preventative survivability is obtained primarily through the application of dedicated protection capacity, while (mesh-based) restoration entails the placement of extra capacity that would be shared by a number of demands and invoked over a variety of failure scenarios. In addition, certain levels of $P()$ may not be feasible as it might require network features that are out of the planner's control, like increasing the connectedness of the underlying network topology. Thus we recommend applications of the model described in the document that allow tighter constraints on $S()$ and more liberal constraints on $P()$.

In planning situations where the demands cannot be determined precisely, planners often rely on conservative over-estimates, thus building networks with some measure of extra capacity. They justify this course of action by pointing out that the additional capital expense is outweighed by the potential for lost revenue and customer defection associated with not being able to satisfy a request for service when it arises. Such networks would be built with the capability to divert capacity to the particular demands that need it as they arise. So it should come as no surprise that the networks naturally have some measure of survivability built-in, especially if the actual demands turn out to be less than what had been planned for. Conversely, many networks that provide a measure of survivability will naturally be robust to some uncertainty in demand. An unexpected increase in some demand could be accommodated by capacity that had been reserved for protection.

A number of schemes that integrate the planning objectives of achieving desired levels of survivability as well as a measure of flexibility to demand uncertainty could be generated, (see, e.g., [9]). The evaluations described in this document would have to be enhanced to include this new dimension. For instance, 100% survivability would mean that every demand *in every scenario* can be restored after any failure. The questions about how to characterize the survivability when some demands may be lost in some scenarios are similarly complicated. An approach to planning for the network may involve explicitly modeling all of the pairs from the demand possibilities and the failure possibilities. This would lead to an extremely large and

intractable mathematical programming problem. An alternative approach would entail developing a network design that is believed to have sufficient flexibility to demands and then, (e.g., by using simulation), determine whether additional capacity should be added to achieve desired levels of survivability. Or a network plan could be developed that consists solely of survivable architectures, like self-healing rings. Then a simulation-based approach to planning under uncertainty would generate variety of feasible demand instances to route. For a given set of capacity values, the formula described in this document would evaluate the amount of survivability provided. Then capacity values could be increased (decreased) so that the objectives regarding survivability can be met.

References

[1] T.H. Wu, *Fiber Network Service Survivability: Architectures, Technologies, and Designs,* Artech House, 1992.

[2] P. Demeester, T.H. Wu, N. Yoshikai, eds. *IEEE Communications Magazine*, Special Issue on Survivable Communication Networks, Vol. 37, August 1999.

[3] M. Mikhail, D. Shallcross, N. Dean, M. Mostrel, "A Commercial Application of Survivable Network Design: ITP/INPLANS CCS Network Topology Analyzer", *Symposium on Discrete Algorithms 1996*, pp. 279-287.

[4] S. Cosares, D. Deutsch. I. Saniee, O. Wasem, "SONET Toolkit: A Decision Support System for the Design of Robust and Cost-Effective Fiber-Optic Networks", *Interfaces* 25, pp. 20-40, 1995.

[5] R. Davis, K. Kumaran, G. Liu, I. Saniee, "SPIDER: A simple and flexible tool for design and provisioning of protected lightpaths in optical networks", *Bell Labs Technical Journal* 6, pp.82-97, 2002.

[6] "RSoft Design Group's MetroWAND Makes Metro Mesh Migration Easy", Press Release, September 2002.

[7] S. Cosares, M. Binbasioglu, "Articulating User Preferences about Survivability in a Network Planning DSS", *Proc. Ninth Americas Conference on Information Systems*, (2003).

[8] Y. Myung, H. Kim, D. Tcha, "Design of Communication Networks with Survivability Constraints", *Management Science*, 45, 238-252, 1999.

[9] J. Kennington, E. Olinick, A. Ortynski, G. Spiride, "Robust Solutions for the WDM Routing and Provisioning Problem with Uncertain Demands and Single Link Failure Protection", Hart eCenter at SMU Technical Report No. HeC-02-01, (2001).