1. INTRODUCTION

Over the past few decades there have been many mathematical models for the design and dimensioning of public and private communication networks. These models have tended to concentrate upon the performance aspects of the network and to determine the link and node capacities which are required to meet prescribed performance objectives of loss and/or delay. The primary objective has been to minimise the cost of (expensive) network plant. However, in recent years with the development of optical fibres having very large capacities available at a relatively cheap cost, the emphasis in network design has shifted from being exclusively performance based, to one in which the users require both performance and reliability/availability as businesses have become more and more dependent on good communications for their effective operation.

Design of networks to meet a set of performance constraints has been the subject of extensive investigations over the past two or three decades in relation to both telephony and, more recently, data networks. Very few models have been developed to jointly consider performance and reliability.

In this paper, we shall attempt to develop mathematical models which incorporate both performance and reliability criteria. A generic mathematical programming model is presented in section 3. Mathematical models for reliability are often quite difficult to solve because of the large solution state space involved. We shall describe an approach for solving this problem for practical size networks using an heuristic method. In section 4 we discuss a number of different realisations of the generic model in the areas of private circuit switched and Asynchronous Transfer Mode networks. Finally, in section 5.1 a simple 8 node network is investigated for a variety of different network modelling scenarios and solutions obtained using the simple heuristic approach which has been developed.

2. NOTATION

We begin our description of the reliability models by establishing the notation to be used.
$L$: is the index set of candidate links.

$R$: is the set of candidate routes for the model.

$\Pi$: is the set of OD pairs for the network.

$S_p$: is the set of candidate routes for OD pair $p$, $p \in \Pi$. $S_p \cup S_q = \emptyset$ for $p \neq q$.

$\lambda_p^r$: is the message/demand arrival rate of the unique OD pair $p$ and associated with route $r$, where $r \in R$.

$\delta_{pl}^r$: is 1 if route $r$ of OD pair $p$ uses link $l$ and 0 otherwise.

$1/\mu$: is the average message length in bits per message.

$n_l$: is the capacity to be installed on link $l$. (This may be expressed in terms of module sizes rather than numbers of circuits in some cases.)

$F_l$: is the average flow in capacity units per second on link $l \in L$.

$C(n_l)$: is the cost of link $l$ with respect to the capacity which has been installed.

$\Delta_{li}$: is the increment of flow on link $l$ due to the failure of link $i$, where $l, i \in L$ and $i \neq l$.

$\beta_l$: is the probability that link $l$ is down (unavailable) – the proportion of time that the link $l$ is not operational.

$\xi_{li}$: is the flow reduction on link $l$ due to the failure of link $i$ and $i \neq l$.

$x_p^r$: is a binary decision variable which is 1 if route $r \in R$ is chosen to carry the flow of its associated OD pair $p$ and 0 otherwise.

$u_p^r$: is a binary decision variable which is 1 if route $r \in R$ is chosen as an alternative route for its OD pair $p$ and 0 otherwise.

In the models which follow, we shall be concerned with determining a pair of link disjoint chains for each OD pair in the network. One chain will be referred to as the primary chain and will represent the route to be used by an OD pair under normal operating conditions. The other chain will be referred to as the secondary chain and this chain will be used as a backup in the event of a single link failure in the primary chain for the OD pair.

3. GENERIC MODEL DESCRIPTION

The generic model described in this section, represents a reworking of the concepts used by Gavish and Neuman in [1] and we have attempted to maintain a consistency of notation between the descriptions.

The objective function for this model can be expressed as:

$$\sum_{l \in L} C(n_l)$$

(1)
where $C(n_l)$ represents the cost of the link $l$ as a function of the link capacity $n_l$. There are many potential performance models which could be adopted here and they could be of either the loss or delay mode of operation according to the type of network being considered. In the following sections, we shall describe simple implementations which incorporate linear, circuit switched and ATM network models.

For the generic model, we follow the approach by Gavish and others in [1] and try to select two link disjoint chains per OD pair such that, if the primary chain should fail, a secondary chain can be selected by the hardware or software to carry the traffic for the OD pair. Note that we do not necessarily require all of the offered traffic to be carried on the secondary chain after failure of the primary chain. Thus, we need to determine the correct link capacities while minimising the network cost function selected above. The decision variables for the problem are the link capacities $n_l$ and the choice of a primary and a secondary chains for carrying the traffic demands of the OD pairs. We let $x_r^p$ and $u_q^p$ be binary-valued decision variables which indicate whether chain $r$ of OD pair $p$ is used as the primary route and chain $q$ is used as the secondary route for this OD pair.

The constraints for the generic model are as follows:

$$\sum_{l \in L} x_r^p u_q^p \delta_r^p \delta_q^p = 0 \quad \forall r, q \in S_p, \quad p \in \Pi$$

Constraint (2) effectively prohibits chains with common links from consideration in the network as well as ensuring that the primary and secondary chains which are selected are also distinct. In other words, they force the network to contain link disjoint paths for the primary and secondary routes. Note that this constraint is actually a quadratic in the decision variables.

$$F_l + \Delta_{li} - \xi_{li} \leq n_l \quad \forall l \neq i \in L.$$  

Constraint (3) ensures that the capacities $n_l$ are not exceeded during a failure of link $i$. (This constraint refers only to the single link failure of link $i$.) The constraint says that under a single link failure of link $i$, the flow on link $l$ which results comprises of flows which are not affected by the failure of link $i$, the increase in flow due to the choice of this link as part of a secondary chain using link $l$ and finally, the reduction in flow on this link as a result of link $i$ forming part of a primary chain which also uses link $l$.

Under normal operating conditions we require:

$$F_l \leq n_l \quad \forall l \in L.$$  

Constraints (3) and (4) also place a lower bound on the value of the unknown capacities $n_l$. The total flow $F_l$ on link $l$ is computed using

$$F_l = \sum_{p \in \Pi} \sum_{r \in S_p} x_r^p \gamma_p$$

where $\gamma_p$ is the traffic demand offered to OD pair $p$. The quantity $\Delta_{li}$ is the flow increment on link $l$ due to the failure of link $i$ and this is computed from:

$$\Delta_{li} = \sum_{p \in \Pi} \sum_{r \in S_p} \sum_{q \in S_p} \gamma_p x_r^p u_q^p \delta_r^p \delta_q^p (1 - \delta_q^p)(1 - \delta_r^p) \quad \forall l \neq i \in L$$
We also note that the OD pairs which use the failed link as part of both their primary and secondary routes will not be able to communicate at all during the breakdown. The flow generated by these OD pairs may be computed from

\[ \Delta_{ll} = \sum_{p \in \Pi} \sum_{r \in S_p} \sum_{q \in S_p, q \neq r} \gamma_p x^p_r u^p_q \delta^p_{ql} \delta^p_{rl} \quad \forall l \in L \]  

(7)

Now, if we restrict ourselves (as do Gavish and Neuman) to the case where we require link disjoint routes, then equation (2) enables equation (7) to simplify to \( \Delta_{ll} = 0 \) and for the case where \( l \neq i \) we find that \( \Delta_{li} \) simplifies to

\[ \Delta_{li} = \sum_{p \in \Pi} \sum_{r \in S_p} \sum_{q \in S_p, q \neq r} \gamma_p x^p_r u^p_q \delta^p_{ri} \delta^p_{ql} \quad \forall l \neq i \in L \]  

(8)

Finally, if both \( l \) and \( i \) belong to the primary route used by various OD pairs, the failure of link \( i \) results in the reduction of the flow which is being carried by link \( l \) since the traffic carried by these OD pairs is shifted to the secondary routes. Hence the decrease in flow on link \( l \) due to the failure of link \( i \) is determined by the flows of those OD pairs \( p \) which satisfy the following requirement:

\[ x^p_r \delta^p_{ri} \delta^p_{ql} = 1 \quad \forall l \neq i \in L \]

thus the quantity \( \xi_{li} \) is given by the expression

\[ \xi_{li} = \sum_{p \in \Pi} \sum_{r \in S_p} \gamma_p x^p_r \delta^p_{ri} \delta^p_{ql} \quad \forall l \neq i \in L \]  

(9)

We also require that

\[ \sum_{r \in S_p} x^p_r = 1, \quad \forall p \in \Pi \]  

(10)

This constraint ensures that one primary route (only) per OD pair is selected as the primary route.

\[ \sum_{q \in S_p} u^p_q = 1, \quad \forall p \in \Pi \]  

(11)

This constraint ensures that one alternative route (only) is chosen for the OD pairs of the network.

\[ x^p_r, u^p_q \in \{0, 1\} \quad \forall r, q \in S_p, \quad p \in \Pi \]  

(12)

These are the decision variables and represent the binary decisions for the chains to be selected for each OD pair in the network.

\[ n_l \geq 0 \quad \text{and integer} \quad l \in L \]  

(13)

The capacities on the links should be non-negative integers.

Thus we see that this model attempts to choose the link capacities \( n_l \) and ensure that there are two link disjoint chains for each OD pair in the network in order to satisfy the
reliability/cost/performance constraints. We also observe that constraint equations (2) and (3) both involve quadratic expressions in the decision variables. However, we find that we may perform a transformation of the quadratic relationships by noting the binary nature of the decision variables. Hence, a constraint from equations (2) of the form

$$x_p^r u_q^r = 0 \quad \text{where } x_p^r, u_q^r \in \{0, 1\}$$

may be replaced with a linear constraint

$$x_p^r + u_q^r \leq 1 \quad \text{where } x_p^r, u_q^r \in \{0, 1\}$$

Thus, if either of the two variables is unity, it forces the other variable to become zero. It also allows both variables to be zero if required. A similar type of transformation is required for constraints (3), since the formula for $\Delta_l$ given in equation (6) can be converted into

$$\Delta_l = \sum_{p \in \Pi} \sum_{r \in S_p} \sum_{q \in S_q} \gamma_p Y_{rq}^p \delta_r^p \delta_q^p \quad \forall l \neq i \in L$$

where $Y_{rq}^p \in \{0, 1\}$ is a binary variable which must satisfy the following conditions:

$$Y_{rq}^p \leq x_r^p$$

$$Y_{rq}^p \leq u_q^p$$

$$x_r^p + u_q^p - 1 \leq Y_{rq}^p$$

Thus we see that $Y_{rq}^p = 1$ whenever both $x_r^p$ and $u_q^p$ are both unity and that $Y_{rq}^p = 0$ in all other cases. Despite the fact that the quadratic nature of the constraints has been removed by this transformation, it is still clear that the problem state space may be quite large and potentially difficult to solve.

It is often said that modern network components are now very reliable and that multiple link failures would be extremely rare. The model also takes this view, however, it should be noted that we could take multiple failures into account if we wished by defining further $\Delta$ and $\xi$ terms which take account of multiple failures.

4. SPECIFIC REALISATIONS OF THE MODEL

4.1. Linear Model

The simplest version of the model involves choosing a cost function which is linear in the link capacities. Such a model could be useful for determining actual or upper bounds on the capacities of links in a private network, or an ATM network, where demands have been specified for the traffic streams in terms of absolute numbers of circuits or bit rates. Thus, the cost function which we can use in such cases is as follows:

$$\sum_{l \in L} \left( \frac{C_l}{1 - \beta_l} \right) n_l$$

where $C_l$ represents the cost per unit demand on link $l$ and the term in the denominator, $1 - \beta_l$, represents the probability that the link is operational. The purpose of the
A probability term is to ensure that, in the event that two chains have equal costs per unit demand, the model will select the most reliable chain.

We have included a number of extensions to the generic model of the previous section. In particular, we have permitted the capacities $n_i$ to be modular. Thus, a user may specify that “circuits” come in appropriately sized bundles for installation in the network. A further enhancement to the model permits the user to specify a “grade of service” which is to apply in the event of a failure affecting the primary chain. This option is included because the secondary chain is (typically) longer and usually more expensive than the primary chain. If we insist that the original demands for end to end capacity are met, then the resulting network may be at least twice the cost of a network which doesn’t take failures into account. This may make the “reliable” network prohibitively expensive to provide. Thus, we have adopted the approach that the end to end demand to flow on the secondary chain can be user specified as a grade of service parameter. The results of applying these options in the linear model are presented in section 5.1.

Solution of the linear cost function model is quite simple to implement. A standard integer linear programming package can be used, or the special structure of the model can be exploited to give a simple algorithm which will actually generate the optimal solution for this case. The details of the basic approach are sketched below:

**Linear Cost Function Algorithm**

1. Generate the costs per unit demand of all chains for an OD pair $p$.

2. Determine the chain with the minimum cost and select it for the primary chain.

3. For each remaining chain, repeat the following procedure until a suitable secondary chain is found:
   (a) Select the next cheapest chain.
   (b) Is this chain link disjoint with respect to the primary chain? If the chain is disjoint, terminate the search as this will be the appropriate secondary chain. Otherwise, eliminate this chain from consideration and choose a new chain.

4. Allocate the demands to each of the primary chains in the network and determine the capacities required to satisfy the demand in terms of circuits and/or modules.

5. Systematically consider each of the links in the network and fail each of the other links. Allocate the demands (adjusted to meet the grade of service standard where required) to each of the links. Maintain the highest required capacity for each link as the allocations proceed.

6. When all links have been considered for single link failures, the process terminates by selecting the highest capacity which was required.
4.2. A Circuit Switched Network Model

The model described in subsection 4.1 was based upon satisfying deterministic demands which do not take into account the economies of scale which are usually present in communications networks. For the circuit switched network model in this subsection, we are interested in private circuit switched networks which do not employ (extensive) alternative routing. (Thus, this model would not be suitable for the Public Switched Telephone Network, but it may be relevant to the case where a company wishes to use its own network under normal operating conditions and to switch to the public network in the event of a failure in the private network.)

In this model, we have used the well-known Berry model formula [2] to determine the number of circuits required on the network links. Once again, we include modularities and the option of specifying different grades of service for the primary and secondary chains. The new cost function is now a nonlinear function of the flows in the network, but we still require that the link capacities must satisfy the normal and failure condition flows. The new cost function now takes into account the stochastic nature of the network flows.

\[
\sum_{l \in L} \left( \frac{C_l}{1 - \beta_l} \right) n_l(F_l) \tag{20}
\]

where we assume that the link capacities \( n_l \) are functions of the flows under normal and failure conditions. We shall assume that the offered traffic demands are Pure Chance and that the designer has specified both normal and failure condition grades of service (end to end performance standards) for the OD pair traffics. The Berry model specifies the number of circuits on a link \( l \) as:

\[
n_l(F_l) = F_l + A_l \left\{ \frac{(M_l - F_l)}{(M_l - F_l - 1)(M_l - F_l) + v_l} - \frac{M_l}{M_l^2 - M_l + V_l} \right\} \tag{21}
\]

where \( M_l \) and \( V_l \) represent the mean and variance of the offered traffic demands, \( v_l \) represents the variance of the traffic overflowing from the link (the traffic is lost in this network scenario) and is given by:

\[
v_l = \frac{(M_l - F_l)}{6} \left\{ 3 - (M_l - F_l) + \sqrt{(3 - M_l + F_l)^2 + 12 A_l \left[ \frac{(M_l - F_l)}{M_l} \right]^{0.7}} \right\} \tag{22}
\]

and \( A_l = V_l + 3 \frac{V_l}{M_l} (\frac{V_l}{M_l} - 1) \) is the well-known Rapp formula for the equivalent random traffic.

4.3. Application to an ATM Network

Finally, we consider a simple implementation of the generic model to an ATM network with bursty traffic streams and based upon a standard Switched Poisson Process (SPP) model. This model accepts OD traffic streams with bursty traffic modelled by the SPP. The simplest version of this model employs the Interrupted Poisson scheme where the peak offered traffic matches the Poisson arrival rate when the switch is on and the IPP parameters are matched to give the average rate of the offered traffic. This model can be used in either a loss or delay mode of operation with the user specifying the appropriate probabilities for the traffic streams. An alternative approach has also been developed for ATM networks and is based on a model by Rogers and Avellaneda [3].
5. EXAMPLES

5.1. Sample 8 Node Network

Let us now consider an 8 node sample network as used by Gavish and Neuman to develop a system of equations to solve the various models considered above. The network is depicted in Figure 1 and has 13 both-way links numbered as shown together with the link costs used.

![Sample 8 node network for model computations](image)

Figure 1. Sample 8 node network for model computations

In this sample network we have chosen to consider the complete set of 56 OD pairs. Chains for each of the OD pairs have been generated using a specially constructed algorithm which has restricted the number of links per chain to a total of 4. This process has produced a total of 408 chains and Table 1 gives the OD pair demands and the distribution of chains between each of the OD pairs. From this table, we are able to deduce the number of variables required for the problem. In this case, there are two variables for each possible chain, since the chain list is required for both the primary and secondary route selection. In addition to these variables we have the capacity variables $n_l$ of which there are 13 in this example. In summary, we require

$$2 \times 408 + 13 = 829$$

variables.

We can also determine the total state space for the problem in terms of the need to identify two separate chains for each OD pair to be selected. This gives an *upper bound*
on the size of the state space as $2.735 \times 10^{72}$ possible states. This number does not take into account the fact that we require the chains to be link disjoint in our final selection for the model, so the state space in this case is actually smaller than this figure suggests!

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>59</td>
<td>45</td>
<td>60</td>
<td>24</td>
<td>33</td>
<td>45</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>-</td>
<td>32</td>
<td>83</td>
<td>43</td>
<td>32</td>
<td>19</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>71</td>
<td>-</td>
<td>23</td>
<td>22</td>
<td>84</td>
<td>36</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>39</td>
<td>48</td>
<td>65</td>
<td>-</td>
<td>15</td>
<td>57</td>
<td>39</td>
<td>59</td>
</tr>
<tr>
<td>5</td>
<td>68</td>
<td>22</td>
<td>88</td>
<td>21</td>
<td>-</td>
<td>16</td>
<td>50</td>
<td>34</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>65</td>
<td>58</td>
<td>50</td>
<td>50</td>
<td>-</td>
<td>32</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>43</td>
<td>91</td>
<td>67</td>
<td>78</td>
<td>50</td>
<td>50</td>
<td>-</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>56</td>
<td>45</td>
<td>22</td>
<td>70</td>
<td>67</td>
<td>50</td>
<td>-</td>
</tr>
</tbody>
</table>

(a) OD Pair Demands

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>9</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>-</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>7</td>
<td>-</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>-</td>
<td>11</td>
<td>6</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>11</td>
<td>-</td>
<td>7</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>-</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
</tbody>
</table>

(b) Total Chains for each OD Pair

Table 1

OD Pair Demands and Chain Information

5.2. Solutions to the 8 Node Problem

Figure 2 summarises the results of applying the linear and circuit switched models to the 8 node network problem. We see that by insisting on very good grades of service the network cost is approximately double the cost of a network in which no provision is made for reliability. The results which have been plotted assume a module size of 1 and the circuit switched model assumes an end to end grade of service of 1%.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance and support for this work received from British Telecom Laboratories, Martlesham Heath, U.K. through a Short-term Fellowship during May - June 1992. The help and encouragement of Dr. Peter Key and the staff from his group is gratefully acknowledged.

REFERENCES

Figure 2. Effects of Varying GOS on the Secondary Chains