Towards the Integration of Reliability and Traffic Engineering

András Faragó
Department of Computer Science
The University of Texas at Dallas
Richardson, Texas, U.S.A.
E-mail: farago@utdallas.edu

***Abstract*—** *The main goal of this paper is to make a step towards the development of a unified approach that incorporates both reliability and traffic engineering aspects into a common, unified model. The approach is motivated by a simple fact: if the traffic flow is blocked between two nodes, then for the end user it is irrelevant why it is happening (whether it is due to failure or congestion). What really matters is that the network should provide sufficient connectivity and bandwidth between the end-nodes, at least with a prescribed probability, taking into account that failure and congestion may both happen. Existing methods do not support such a joint analysis of network reliability and traffic, since the two scenarios are addressed by different models. The proposed approach can provide a tight and unified estimate of the probability of any event in the network that is expressible via the availability of sets of links. It takes both reliability and traffic engineering aspects into account in an algorithmically efficient way.*

Keywords: *Traffic engineering, blocking probability, network reliability, teletraffic theory, computational complexity.*

I. Introduction

In traffic engineering and network design it is a task of fundamental importance to compute or estimate the probability that between a source-destination pair of nodes at least one route is available with sufficient capacity for carrying traffic flow. This task is a key part of many traffic engineering and network design procedures.

There are two typical reasons for the non-availability of a given route. The *reliability based* reason is that at least one link or node is inoperational along the route. The *traffic based* reason is that at least one link or node is congested (blocked) on the route. The reliability and traffic based scenarios are typically analyzed independently of each other, since different models handle the two aspects. On the other hand, from the end user point of view, it may not matter *why* it is impossible to send the traffic flow (whether it is due to failure or congestion). What really matters is that the network should be able to provide an available route with sufficient bandwidth between the end-nodes, at least with a prescribed probability.

Why is it difficult to compute the probability that at least one route is available between the source and destination? As explained below, each of the reliability and traffic aspects alone already creates a very hard algorithmic problem. Naturally, if we want to consider them jointly, no relief from the hardness can be expected.

The *reliability models* address only one part of the problem, as they do not take into account traffic-dependent blocking probability. Nevertheless, even in the purest reliability model, it is well known that all the relevant reliability computations are #*P*-complete [1] where #*P* is a complexity class that is deemed even harder than the well known class *NP*. In #*P* problems we want to know the exact *number* of solutions, as opposed to *NP* problems, where we ask only whether a solution *exists*. Consequently, #*P*-complete problems are algorithmically at least as hard as *NP*-complete ones. (Actually, it is conjectured that #*P*-complete problems are even harder, in the sense that even if we could solve an *NP*-complete problem efficiently, it would still not provide a way to also solve #*P*-complete problems efficiently.) In any case, the *exact* reliability computation, even with ignored traffic aspects, is algorithmically very hard.

In the *traffic engineering models* the possibility of traffic-independent failure (i.e., the reliability aspect) is ignored. The most well-known network level model for computing route blocking probabilities is the *reduced load approximation* of Kelly and Whitt [4]. In this model the load of a given link is computed by taking into account that the routes which pour their traffic in the link suffer some loss on other links, too. This reduces the load of the considered link, hence the name reduced load approximation. The reduction is computed using the *link independence assumption*, i.e., assuming that the event of blocking is independent on different links. This independence assumption, however, can be questioned in a number of cases. An example is when lightpaths are set up in WDM networks with no wavelength conversion. In this case, the wavelength continuity requirement can create a situation that two adjacent links are both have available capacity, yet there is no available route that goes through both, due to the lack of a common available wavelength. Clearly, this invalidates the link independence assumption. Additionally, even if link independence is assumed, the reduced load approximation does not support reliability considerations. The picture is further complicated by the fact that link independence often does not hold form the reliability point of view, either. For example, when overlay logical neworks exist on top of a physical network, then the failure of a physical link disconnects all logical links carried by the physical link. Therefore, these logical links do not fail independently.

The reduced load approximation was originally developed for the classical case of single-service Poisson traffic, but it can be extended to more general cases, as well, but still assuming link independence [4]. It is also known that the model can be made exact, but that results in an exponentially growing set of equations [2].

As a general exact approach, of course, one can always consider a detailed Markov model that takes into account all possible states of the network. This is, however, very far from being scalable, due to state space explosion. Even if it is restricted only to traffic (with no reliability considerations), and the simplest classical Erlang traffic model is used, it still involves algorithmically very difficult tasks. For example, it is known that computing the route blocking probabilities in an exact Markov model is #*P*-complete [3], and, consequently, it is at least as hard as any *NP*-complete problem (but actually conjectured even harder).

Scalability can be significantly improved if we are satisfied with bounds/approximations instead of the exact values. Various approximations are known both for traffic and reliability, but each addresses only one of the two aspects.

The goal of this paper is to make a step towards developing an integrated approach to handle reliability and traffic engineering issues in a common, unified model. Specifically, we solve the following core problem. Given a network with a set of routes, each with known offered end-to-end traffic. Each link can fail with a given probability and can also block due to congestion. The blocking probability is a general function of the link load, the approach is not restricted to a specific traffic model. The link loads are not known in advance, as they are the results of the route-carried traffic volumes that depend on the blocking and failure of other links. Moreover, these events are not necessarily independent. Under these conditions we provide upper and lower bounds for the probability that at least one of k given routes is available between a source-destination pair. In fact, the approach is more general and can evaluate the probability that at least one set of links among of k arbitrary sets of links is available. All the provided bounds are computable efficiently (in polynomial time) for any fixed value of k and they can be coupled with any traffic model. We also show that the bounds can be iteratively refined, yielding a polynomial-time approximation scheme for computing the route blocking probabilites.

II. The Model

Let us denote the links of the network by $1, \ldots, m$. Links are considered one-directional, i.e., if a physical link between the same nodes is used in both directions then in the model this is represented as two directed links. We are also given a set $\mathcal{R}$ of traffic carrying routes in the network, where each route $r \in \mathcal{R}$ is a nonempty set of links (which does not need to be a simple path in the graph theoretic sense). The incidence of links and routes will be most conveniently described for our purposes by introducing the sets $\mathcal{R}_j \subseteq \mathcal{R},\ j = 1, \ldots, m$, where $\mathcal{R}_j$ is defined as the set of routes that contain link j. It is assumed that each link is used by at least one route, that is, $\mathcal{R}_j \neq \emptyset$.

For each route $r \in \mathcal{R}$ the *end-to-end offered traffic* is denoted by ν_r. We assume a stationary, but otherwise arbitrary arrival process with an arbitrary holding time distribution for which the expected value exists. For a single service system ν_r is the arrival rate multiplied by the average holding time. For multiservice systems ν_r is a vector in which the coordinates are the offered traffic values in the respective traffic classes.

The offered traffic (vector in the multiservice case) to link j will be denoted by ρ_j. As opposed to ν_r, however, the value of ρ_j is not an input parameter, that is, its value is not known in advance. The link-offered traffic ρ_j will result as a superposition of the route-*carried* traffic volumes, which are influenced by the availability of the links that may in turn depend on each other, thus creating a complex mesh of dependencies.

Consider now a link j. Let p_j be the reliability of link j, that is, the link is operational with probability p_j and it is down with probability $1 - p_j$. The link also has an offered traffic of ρ_j (which, as mentioned in the previous paragraph, is not known in advance). The individual behavior of link j is described by a *link availability function* $\alpha_j(p_j, \rho_j)$. This function is based on a reliability and traffic model and maps the reliability p_j and offered traffic ρ_j into the probability of availability of the link. This is the probability that the link can serve a request, i.e., it is up and not blocked. The approximative assumption we make here is that the link behavior depends on the *aggregate* traffic load of the link, independently of how it is composed from individual traffic streams. On the other hand, we do not assume that the blocking or unavailability of different links is necessarily independent.

An essential point is that the *structure* of the model and of the solution does not depend on the actual form of the function $\alpha_j(p_j, \rho_j)$. Thus, $\alpha_j(p_j, \rho_j)$ can be any function of its parameters, with only two very natural requirements:

- *(i)* $0 \le \alpha_j(p_j, \rho_j) \le 1$, since $\alpha_j(p_j, \rho_j)$ represents a probability.
- *(ii)* $\alpha_j(p_j, \rho_j)$ is a decreasing function of ρ_j. This means that higher traffic load can only decrease the probability that the link is available.

In this way we can achieve that the link availability function becomes essentially a *replaceable subroutine* in the network level model, so we do not depend on a specific reliability and traffic modeling approach. Thus, $\alpha_j(p_j, \rho_j)$ can be any function, determined by "plugging in" any specific reliability and traffic model. A simple example is when traffic is captured by the classical single service Erlang model and the reliability of the link is independent of the blocking probability. In this case we have

$$\alpha_j(p_j, \rho_j) = p_j\big(1 - E_n(\rho_j)\big) = p_j\left(1 - \frac{\rho_j^n/n!}{\sum_{i=0}^{n}\rho_j^i/i!}\right)$$

where n is the capacity of the link (number of circuits/channels) and $E_n(\rho)$ is the well-known Erlang B formula.

One can, of course, have more sophisticated cases, as well. For example, let us assume that multi-service traffic is present (in this case ρ_j is a vector) and let $B_j(\rho_j)$ be the blocking probability of link j, computed by some complex multi-service traffic model, such as the ones found in [4]. Further, let the link reliability also depend on the traffic load. For example, let $p_j = \frac{a_j}{b_j + ||\rho_j||}$, where a_j, b_j are constants and $||.||$ is some vector norm. The rationale behind traffic dependent reliability is that the network may be a tactical network with wireless links, operating under hostile conditions, and the links with higher traffic tend to have higher chance to get jammed or attacked. In this case a possible link availability function is

$$\alpha_j(p_j, \rho_j) = \frac{a_j}{b_j + ||\rho_j||}\big(1 - B_j(\rho_j)\big).$$

Rather than assuming, however, this or another specific link model, we intentionally allow any general function for $\alpha_j(p_j, \rho_j)$, subject to the natural conditions mentioned under *(i)* and *(ii)*.

Another important point is that the link availability function describes how the link behaves *in isolation*. When the links are connected into a network, it is quite possible that the availability of different links is *not* independent. For example, if links i and j have reliabilities p_i, p_j and offered loads ρ_i and ρ_j, respectively, then they are individually available with probabilities $\alpha_i(p_i, \rho_i)$ and $\alpha_j(p_j, \rho_j)$. On the other hand, due to correlated link failures, it may well be the case that

$$\mathrm{P}(i, j \text{ both available}) \neq \alpha_i(p_i, \rho_i)\alpha_j(p_j, \rho_j).$$

For a link j let $p(j|r)$ denote the *conditional reliability* of the link, under the condition that the links of a given set r of links do not fail. With this notation we can write, using $r = \{j\}$ for the above example,

$$\mathrm{P}(i, j \text{ both available}) = \alpha_i(p(i|\{j\}), \rho_i)\alpha_j(p_j, \rho_j).$$

In case of more than two links the expression would become more complicated. Nevertheless, if a failure model is fixed, then one can, in principle, always compute $p(j|r)$, by summing the probability of the appropriate states. This, of course, may have high complexity. On the other hand, note that in the most frequently used independent failure model simply $p(j|r) = p_j$ holds. To keep our treatment simple, in the rest of the paper we use the reliability model that assumes independent link failures. In this way the formulation of our results become simpler, yet still shows the essential ideas. All the results, however, can be directly generalized for more general failure models, by replacing the reliability values by the conditional ones.

A route is called available if each of its links are simultaneously available, that is, they are all up and not blocked. Due to potential link dependence, as mentioned above, the probability of route availability may not be computed simply as the product of link availability probabilities. (Note that even if the independent failure model is used for reliability, traffic blocking is not assumed to be necessarily independent.)

Given the above input, we are primarily interested in finding efficient (polynomial-time computable) upper and lower bounds on the probability that at least one of the routes among k given routes are available (while taking into account the overall network traffic and potential failures).

III. Lower and Upper Bounds on the Route Availability Probabilites

Let us introduce a few notations. For any set r of links let $A(r)$ denote the (so far unknown) probability that the whole set is available, in the sense that each link in the set is simultaneously available. If $r = \{j\}$, i.e., r consists of a single link, then $A(r) = \alpha_j(p_j, \rho_j)$. In most applications r is a route, so we refer to $A(r)$ as route availability probability. Nevertheless, in general, r can be any set of links. To avoid degenerated cases we always assume $A(r) > 0$, i.e., there is no route (or link set) that is *never* available. Recall that we drop the link independence assumption, so, in general, $A(r) \neq \prod_{j\in r} \alpha_j(p_j, \rho_j)$ may hold.

We use the notation $r^{[\ell]}$ to denote the *initial segment of length* ℓ of route r, or, in general, the subset of r that consists of the first ℓ links of r, for $\ell \leq |r|$. The initial segment, of course, may depend on how the links are numbered, but we simply assume any fixed numbering $1, \dots, m$ of all links in the network. If $\ell = 0$ then $r^{[0]} = \emptyset$, i.e., the initial segment of length 0 is the empty set for which we define $A(\emptyset) = 1$.

Now we are ready to define two fundamental quantities that will be shown to be lower and upper bounds on route availability. For any set $r = \{j_1, \dots, j_s\}$ of links, we define these bounds as follows, using the introduced notations:

$$A^-(r) = \prod_{\ell=1}^{s} \alpha_{j_\ell} \left(p_{j_\ell}, \sum_{q\in\mathcal{R}_{j_\ell}} \nu_q \right) \qquad (1)$$

and

$$A^+(r) = \prod_{\ell=1}^{s} \alpha_{j_\ell} \left(p_{j_\ell}, \sum_{q\in\mathcal{R}_{j_\ell}} \nu_q A^-\big(q - r^{[\ell]}\big) \right). \qquad (2)$$

Now we are ready to prove the first result.

Theorem 1. *The availability probability $A(r)$ of any set r of links always satisfies*

$$A^-(r) \leq A(r) \leq A^+(r)$$

where A^-, A^+ are defined in (1) *and* (2). *Moreover, if the link availability functions $\alpha_j(p_j, \rho_j)$ can be computed in polynomial time, then both bounds can also be computed in polynomial time.*

Proof. We first show that the route availability probabilities always satisfy the following system of equations:

$$A(r) = \prod_{\ell=1}^{s} \alpha_{j_\ell} \left(p_{j_\ell}, \sum_{q\in\mathcal{R}_{j_\ell}} \nu_q A\big(q - r^{[\ell]}\big) \right); \quad \forall r \quad (3)$$

where j_ℓ, $\ell = 1, \dots, s$, are the links of r.

To prove (3) take a set r of links with $r = \{j_1, \dots, j_s\}$. If $s = 1$ then $r = \{j_1\}$ and (3) becomes

$$A\big(\{j_1\}\big) = \alpha_{j_1} \left(p_{j_1}, \sum_{q\in\mathcal{R}_{j_1}} \nu_q A\big(q - \{j_1\}\big) \right). \qquad (4)$$

In this expression the second argument of $\alpha_{j_1}(.,.)$ is the sum of the offered traffic on all routes that traverse link j_1, each thinned by a factor $A\big(q-\{j_1\}\big)$ for the respective route $q \in \mathcal{R}_{j_1}$, which is, by definition, the probability that the rest of route q is available. Therefore, the result of the summation in the second argument of $\alpha_{j_1}(.,.)$ is the offered traffic ρ_{j_1} to link j_1. Since, by definition, $A(\{j_1\}) = \alpha_{j_1}(p_{j_1}, \rho_{j_1})$ holds, therefore, (3) is valid for $s = 1$.

To handle the case $s > 1$, let us introduce the events a_ℓ, $\ell = 1, \dots, s$, where a_ℓ denotes the event that link j_ℓ is available. Then we can write

$$\begin{aligned} A(r) &= \mathrm{P}(a_1 \cdots a_s) \\ &= P(a_1) \cdot \frac{\mathrm{P}(a_1 a_2)}{\mathrm{P}(a_1)} \cdots \frac{\mathrm{P}(a_1 a_2 \cdots a_s)}{\mathrm{P}(a_1 \cdots a_{s-1})} \\ &= \mathrm{P}(a_1) \prod_{\ell=2}^{s} \mathrm{P}(a_\ell | a_{\ell-1} \cdots a_1). \end{aligned} \qquad (5)$$

Note that this holds regardless to whether the events a_ℓ, $\ell = 1, \dots, s$, are independent or not. The conditional probability $\mathrm{P}(a_\ell | a_{\ell-1} \cdots a_1)$ is the probability that link

j_ℓ is available, given that links $j_{\ell-1},\dots,j_1$ are available. This can be computed by taking into account that now, by the condition, the route-offered traffic for each route q is thinned only by the availability probability of the remaining subset of the route. This subset is $q-\{j_\ell\}-\{j_{\ell-1},\dots,j_1\}=q-r^{[\ell]}$, so we have

$$\mathrm{P}(a_\ell|a_{\ell-1}\cdots a_1)=\alpha_{j_\ell}\left(p_{j_\ell},\sum_{q\in\mathcal{R}_{j_\ell}}\nu_q A\big(q-r^{[\ell]}\big)\right). \tag{6}$$

Now if we take into account $\mathrm{P}(a_1)=A(\{j_1\})$, then the substitution of (4) and (6) into (5) yields precisely the desired system (3).

We know that $A\big(q-r^{[\ell]}\big)\le 1$, being a probability. Therefore, removing it from the righthand-side of (3) can only increase the second argument of $\alpha_{j_\ell}(.,.)$ in (3). As the link availability functions are assumed decreasing in the second argument (see *(ii)* in the Model section), this removal can only decrease the righthand-side of (3). Thus, we obtain

$$A(r)\ge\prod_{\ell=1}^{s}\alpha_{j_\ell}\left(p_{j_\ell},\sum_{q\in\mathcal{R}_{j_\ell}}\nu_q\right)$$

which, by (1), is the same as $A(r)\ge A^-(r)$.

Let us now replace $A\big(q-r^{[\ell]}\big)$ in the righthand-side of (3) by $A^-\big(q-r^{[\ell]}\big)$. Since we already know that $A(r)\ge A^-(r)$ holds for every r, therefore, this replacement can only decrease the second argument of $\alpha_{j_\ell}(.,.)$ in (3). Again, by the decreasing nature of the link availability function, the righthand-side of (3) can only grow this way, yielding

$$A(r)\le\prod_{\ell=1}^{s}\alpha_{j_\ell}\left(p_{j_\ell},\sum_{q\in\mathcal{R}_{j_\ell}}\nu_q A^-\big(q-r^{[\ell]}\big)\right).$$

which, taking (2) into account, is equivalent to $A(r)\le A^+(r)$.

Regarding the complexity of computing the bounds, let m be the number of links and set $R=|\mathcal{R}|$. (Note that the routes along with the traffic is sent are listed as part of the input.) Since for any link j it holds that $\mathcal{R}_j\subseteq\mathcal{R}$, therefore, $|\mathcal{R}_j|\le R$. This implies that the summation on the righthand-side of (1) can be computed in $O(R)$ time. As $|r|\le m$, we have that by calling the link availability function at most $O(m)$ times and doing an $O(R)$ amount of side-computation for each call, we can compute $A^-(r)$. Consequently, $A^-(r)$ can be computed in polynomial time, given that the link availability functions are computable in polynomial time.

In the computation of $A^+(r)$, according to (2), we operate similarly, but with the added complexity that for each ℓ and ν_q we also have to compute the value of $A^-\big(q-r^{[\ell]}\big)$. Since, however, we already know that $A^-(r)$ is computable in polynomial time for any argument r, therefore, it is clear that the added complexity does not destroy the polynomial-time computability of $A^+(r)$. □

IV. Bounds for Sets of Routes

Let us now consider *sets of routes*, rather than just a single route. This makes it possible to answer questions like "what is the probability that at least one of k given routes between a source-destination pair is availabale?"

We need to define some quantities that will serve in the bounds for sets of routes. Let $r_1,\dots,r_k$ be k routes (in general, k sets of links). Using the above introduced functions $A^-(r),A^+(r)$, define for $i=1,\dots,k$

$$S_i^-(r_1,\dots,r_k)=\sum_{j_1<\ldots<j_i}A^-(r_{j_1}\cup\ldots\cup r_{j_i}), \tag{7}$$

that is, the summation runs over all possible unions of i sets out of $r_1,\dots,r_k$, without repetition. Similarly, define

$$S_i^+(r_1,\dots,r_k)=\sum_{j_1<\ldots<j_i}A^+(r_{j_1}\cup\ldots\cup r_{j_i}) \tag{8}$$

If no ambiguity arises, we simply write S_i^- and S_i^+ instead of $S_i^-(r_1,\dots,r_k)$ and $S_i^+(r_1,\dots,r_k)$, respectively.

Finally, for any k sets $r_1,\dots,r_k$ of links let

$$A(r_1\vee\ldots\vee r_k)$$

denote the probability the *at least* one of the sets is fully available (i.e., each link in the set is available). In general $A(r_1\vee\ldots\vee r_k)\ne A(r_1)+\ldots+A(r_k)$, since the availability of $r_1,\dots,r_k$ are not mutually exclusive events. Similarly, let $A(r_1\wedge\ldots\wedge r_k)$ be the probability that they are *all* fully available. Note that, due to not assuming independence in our model, it may be that $A(r_1\wedge\ldots\wedge r_k)\ne A(r_1)\cdot\ldots\cdot A(r_k)$. On the other hand, $A(r_1\wedge\ldots\wedge r_k)=A(r_1\cup\ldots\cup r_k)$, since the simultaneous availability of $r_1,\dots,r_k$ means that all links in the k sets must be available, which is equivalent to the availability of the link set $r_1\cup\ldots\cup r_k$.

Now we can state bounds for k link sets.

Theorem 2 *For any k sets $r_1,\dots,r_k$ of links the following lower and upper bounds hold for the probability* *that at least one of the sets is fully available:*

$$\sum_{i=1}^{\lceil k/2\rceil} S_{2i-1}^{-} - \sum_{i=1}^{\lfloor k/2\rfloor} S_{2i}^{+} \leq A(r_1 \vee \ldots \vee r_k)$$

and

$$A(r_1 \vee \ldots \vee r_k) \leq \sum_{i=1}^{\lceil k/2\rceil} S_{2i-1}^{+} - \sum_{i=1}^{\lfloor k/2\rfloor} S_{2i}^{-}$$

where S_i^-, S_i^+ *are defined in* (7) *and* (8)*. Moreover, if the link availability functions* $\alpha_j(p_j, \rho_j)$ *can be computed in polynomial time, then both bounds can be computed in polynomial time for any fixed* k*.*

Proof. Let us now consider k sets of links, $r_1 \ldots, r_k$. By the well-known inclusion-exclusion principle of elementary probability theory we have

$$A(r_1 \vee \ldots \vee r_k) = \sum_{i=1}^{k} (-1)^{i+1} \sum_{j_1 < \ldots < j_i} A(r_{j_1} \wedge \ldots \wedge r_{j_i}).$$

Due to $A(r_{j_1} \wedge \ldots \wedge r_{j_i}) = A(r_{j_1} \cup \ldots \cup r_{j_i})$, this translates to

$$A(r_1 \vee \ldots \vee r_k) = \sum_{i=1}^{k} (-1)^{i+1} \sum_{j_1 < \ldots < j_i} A(r_{j_1} \cup \ldots \cup r_{j_i}). \tag{9}$$

Since we already know $A^-(r) \leq A(r) \leq A^+(r)$, we can decrease the righthand-side by replacing A by A^- in the terms of positive sign and by A^+ in the terms of negative sign. In this way we obtain

$$\begin{aligned} A(r_1 \vee \ldots \vee r_k) \quad \geq \quad & \sum_{i=1}^{\lceil k/2\rceil} \sum_{j_1 < \ldots < j_i} A^-(r_{j_1} \cup \ldots \cup r_{j_i}) \\ & - \sum_{i=1}^{\lfloor k/2\rfloor} \sum_{j_1 < \ldots < j_i} A^+(r_{j_1} \cup \ldots \cup r_{j_i}) \end{aligned}$$

which is, by the definition of S_i^-, S_i^+, is the same as

$$\sum_{i=1}^{\lceil k/2\rceil} S_{2i-1}^{-} - \sum_{i=1}^{\lfloor k/2\rfloor} S_{2i}^{+} \leq A(r_1 \vee \ldots \vee r_k),$$

thus proving the lower bound in Theorem 1. Similarly, by replacing A by A^+ in the terms of positive sign and by A^- in the terms of negative sign in (9), we obtain

$$\begin{aligned} A(r_1 \vee \ldots \vee r_k) \quad \leq \quad & \sum_{i=1}^{\lceil k/2\rceil} \sum_{j_1 < \ldots < j_i} A^+(r_{j_1} \cup \ldots \cup r_{j_i}) \\ & - \sum_{i=1}^{\lfloor k/2\rfloor} \sum_{j_1 < \ldots < j_i} A^-(r_{j_1} \cup \ldots \cup r_{j_i}) \end{aligned}$$

which is equivalent to

$$A(r_1 \vee \ldots \vee r_k) \leq \sum_{i=1}^{\lceil k/2\rceil} S_{2i-1}^{+} - \sum_{i=1}^{\lfloor k/2\rfloor} S_{2i}^{-}$$

proving the upper bound.

Regarding the complexity of the computation, oberve that for *constant* k both S_i^- and S_i^+ contain only a constant number of terms, each of the form of $A^-(r)$ or $A^+(r)$ for some link set r. Then the polynomial running time follows from the polynomial-time computability of $A^-(r)$ and $A^+(r)$ that is known from Theorem 1. ☐

It is worth considering the important special case $k = 2$, when the bounds become much simpler. In this case we are interested in the probability that at least one of two link-sets is fully available. For example, at least one of two given routes is available (under network traffic as well as potential failures). Substituting $k = 2$ in the general formulas we can obtain the following:

Corollary 1. *For any two sets* r_1, r_2 *of links*

$$A^-(r_1) + A^-(r_2) - A^+(r_1 \cup r_2) \leq A(r_1 \vee r_2)$$

and

$$A(r_1 \vee r_2) \leq A^+(r_1) + A^+(r_2) - A^-(r_1 \cup r_2)$$

hold, where A^-, A^+ *are defined in* (1) *and* (2)*. Both bounds can be computed in polynomial time, given that link availability functions* $\alpha_j(p_j, \rho_j)$ *can be computed in polynomial time.*

Note that the above bounds depend on both the offered traffic values ν_r and on the link reliabilities, as well. This dependence is through the link availability functions $\alpha_j(p_j, \rho_j)$ that are used in the functions A^-, A^+, which in turn are used in the definition of S_i^-, S_i^+. Note that the ρ_j values (link-offered traffic) are not known as input, they are the results of pouring route-offered traffic in the links, thinned by the loss due to blocking or failure of other links. On the other hand, the estimations remain valid for *any* choice of the link availability functions $\alpha_j(p_j, \rho_j)$, given that they satisfy the natural conditions *(i), (ii)* described in the previous section. Thus, we might say, the "beauty" of the approach is that the dependence on the individual link behavior (and, therefore, on the traffic model) is localized as a replaceable subroutine, not affecting the general stucture of how the network level model is built on top of the link models. This makes it possible to incorporate the traffic and reliability considerations in a common, unified model. This is contrast to previous approaches that are built on a specific traffic and link model.

It is worth noting that the bounding technique presented in this section can be extended to provide lower and upper bounds for the probability of *any* event that is expressible with the route blocking events. Due to space limitations we omit the detailed description of this generalization.

V. Iterative Refinement of the Bounds

One can observe that the quantities $A^{-}(r), A^{+}(r)$, defined by (1), (2), respectively, are just the first two nontrivial members $A_1(r), A_2(r)$ of the following iterated sequence:

$$A_0(r) = 1 \tag{10}$$

$$A_{k+1}(r) = \prod_{\ell=1}^{s} \alpha_{j_\ell} \left(p_{j_\ell}, \sum_{q \in \mathcal{R}_{j_\ell}} \nu_q A_k\big(q - r^{[\ell]}\big) \right)$$

It follows from the definitions and from the decreasing property of the α_j functions (see Section II) that

$$A_1(r) \le A_3(r) \le A_5(r) \le \ldots \le A(r) \tag{11}$$

and

$$A(r) \le \ldots \le A_4(r) \le A_2(r) \le A_0(r) \tag{12}$$

hold. In other words, the iterated values $A_0(r), A_1(r), A_2(r), \ldots$ provide more and more tight upper and lower bounds for the exact route availability probability $A(r)$, alternatingly bounding $A(r)$ from above and below.

It is natural to ask whether this iteration converges to the *exact* availability value $A(r)$. It is clear that the sequence of upper bounds is convergent, since it is decreasing and bounded from below. Similary, the sequence of lower bounds is increasing and bounded from above, so it also converges. Without any further condition, however, it may be possible that the two limits do not coincide. For example, if the α_j functions are discontinuous then the iteration may infinitely oscillate between the two sides of a discontinuity. We conjecture, however, that for continuous link availability functions the two limits coincide, thus the iteration converges to the exact route availability $A(r)$. At this time, this problem is still open.

VI. Conclusion

In this paper our main goal was to make the a step towards the development of a unified approach that incorporates both reliability and traffic engineering aspects into a common, unified model. We have proved upper and lower bounds for route availability probabilites in a general network model that allows the incorporation of both traffic and reliability aspects. Since the routes in our model can be arbitrary link sets, using the upper and lower bounds, we can estimate the availability of arbitrary events that are expressible with the availability of link sets.

References

[1] C.J. Colbourn, *The Combinatorics of Network Reliability*, Oxford Univ. Press, 1987.

[2] A. Faragó, "A General Method for the Blocking Analysis of Networks with Dependent Links", *2001 IEEE Workshop on High Performance Switching and Routing*, Dallas, Texas, May 29–31, 2001, pp. 124–129.

[3] G. Louth, M. Mitzenmacher and F.P. Kelly, "Computational Complexity of Loss Networks", *Theoretical Computer Science*, 125(1994), pp. 45-59.

[4] K.W.Ross, *Multiservice Loss Models for Broadband Telecommunication Networks,* Springer, 1995.