End-To-End Reliability-Dependent Pricing of Network Services

Bruno TUFFIN[1], **Pablo RODRÍGUEZ BOCCA**[2] **and Héctor CANCELA BOSI**[2 *]
[1] IRISA-INRIA, Rennes, France
[2] Departamento de Investigación Operativa, Instituto de Computación, Facultad de Ingeniería,
Universidad de la República, Montevideo, Uruguay
btuffin@irisa.fr, prbocca@fing.edu.uy, cancela@fing.edu.uy

Abstract:

This work is a first step in the direction of charging telecommunication networks access based on the reliability of end-to-end paths. In the literature, congestion pricing is usually proposed to guarantee quality-of-service requirements of Internet applications like multimedia transmission. Nevertheless, with the widespread deployment of optic-fiber, some authors believe that capacity will be ahead of demand, and this pricing mechanism will not be applied. We note that even with infinite link capacity, availability is still a concern. We then present a charging scheme based on reliability, which is used to fix optimal prices and also to extend an existing network.

Keywords: *network design, pricing schemes, reliability, metaheuristics, genetic algorithms.*

1 Introduction

Devising new charging schemes for telecommunication networks has become a hot topic in the scientific community, as the current flat-rate charges used in the Internet are an incentive for overusing the resources and traffic continues to grow exponentially, which may result in congestion. Also, the network has to deal with applications having different quality of service (QoS) requirements. For instance, voice and video over IP require small delays and jitter, but can support some losses, whereas e-mail or file transfers do not support losses but are not delay sensitive. To ensure QoS for the different applications, a service differentiation scheme has to be devised, like in IntServ or Diffserv architectures. A pricing scheme has to be attached to it, otherwise all customers will choose the best available service class, and the service differentiation becomes meaningless. Some exhaustive surveys are the works by Da Silva (2000), Falkner et al (2000), Henderson (2001), and Tuffin (2003).

One alternative viewpoint that we consider here is to observe that with the widespread use of optic fiber and improving technologies, the backbone networks are and probably will be over-dimensioned, so that in general congestion will not occur. In this context, it could be more suitable to charge the network access, based on connection reliability.

We consider a network topology, where each link is assumed to have an infinite capacity (corresponding to over-dimensioning) but may fail with a given probability. Each pair of nodes has then a probability to be connected. The price for each connection between a source *s* and a destination *t* depends on the reliability of the connection between *s* and *t*. Of course the demand is also varying with this price, so that a first goal, discussed in Section 2, is to set up a price that maximizes the network revenue. In a second stage, the problem is to extend or in general modify the topology of an existing network in order to increase the service provider's net profit (the revenue minus the cost of modifying the network); this problem is formulated in Section 3. In Section 4 we describe a genetic algorithm which is used to find an approximate solution; and

---

[*] corresponding author tel: +598-2-7114244 ext. 113; fax: +598-2-7110469

weapply it to solve a problem whose topology is inspired by the VTHD (Very High Broadband IP/WDM test platform) French network.

2 Pricing model

We consider the network of an Internet Service Provider (ISP), represented by an undirected communication network $G = (N,E)$ consisting of a set of nodes N and a set of connecting links E. Let m be the number of links and n the number of nodes of G. We assume that each link will be over-provisioned, so that each link has effectively infinite capacity. We consider that for each edge $l \in E$, we can choose between different technology types, which have different cost and probability of failure. For each link $l \in E$, $T(l) \subseteq T$ is the set of all the possible technology types applicable to link l. So, for each link $l \in E$ and each technology type of this link $t \in T(l)$, we assign an (independent from others) probability of failure $q_l(t)$ and a cost $c_l(t)$ (which can depend on the link length, the geography, technology amortization, operation and management associated costs, etc.). For simplicity, we assume that nodes do not have cost and that they do not fail. Only a subset of nodes $K \subseteq N$ has connection demands, we call these nodes *terminals*. To each pair of terminals (s,t) with $s,t \in K$, we associate a total connection demand rate $\tilde{\lambda}_{s,t}$, a connection duration (assumed to be exponential with rate $\mu_{s,t}$), and a reliability $r_{s,t}$, which correspond to the probability that nodes s and t are connected. Arrivals and connection durations are assumed to be independent. To each pair of nodes (s,t) is associated the utility function of getting a connectionwith reliability r, modeled by a random variable $U_{s,t}(r)$, expressed in monetary units. The overall level of satisfaction is then $U_{s,t}(r)-p$ where p is the connection price. A customer will enter the network if and only if $U_{s,t}(r) \geq p$. The random variable $U_{s,t}(r)$ is characterized by is distribution: we denote by $F_{\mathrm{r}s,t}$ its cumulative distribution function and we define $\overline{F}_{r_{s,t}} = 1 - F_{r_{s,t}}$. Our goal is then to find out the optimal prices, for each pair (s,t), in terms of the reliability $r_{s,t}$, maximizing the network revenue $H(G) = \sum_{(s,t)\in K} n_{s,t} p_{s,t}$ over the set of prices $p_{s,t} \geq 0$, for all s,t, with $n_{s,t}$ mean number of online (s,t)-connections.

The actual arrival rate of connections between s and t, $\lambda_{s,t}(p_{s,t})$, is given by $\lambda_{s,t}(p_{s,t}) = \tilde{\lambda}_{s,t} P(U_{s,t}(r_{s,t}) \geq p_{s,t}) = \tilde{\lambda}_{s,t} \overline{F}_{r_{\overline{F}_{r_{s,t}}}}(p_{s,t})$. Also, according to classical queueing theory for the M/M/∞ queue, we have $n_{s,t} = \frac{\lambda_{s,t}}{\mu_{s,t}}$ so that $H(G) = \sum_{(s,t)\in K} \frac{\tilde{\lambda}_{s,t}}{\mu_{s,t}} p_{s,t} \overline{F}_{r_{\overline{F}_{r_{s,t}}}}(p_{s,t})$.

If we use first order conditions while maximizing this revenue (as the price $p_{s,t}$ is necessarily positive otherwise the revenue between s and t would be zero, meaning that the Langrange multiplier is zero), i.e., $\partial G/\partial p_{s,t}=0$. s,t, (assuming that it gives the solution) we get $\frac{\partial}{\partial p_{s,t}}\left(p_{s,t}\overline{F}_{s,t}(p_{s,t})\right) = 0$, that is $\overline{F}_{s,t}(p_{s,t}) + \frac{\partial \overline{F}_{s,t}(p_{s,t})}{\partial p_{s,t}} = 0$

In general, solving these equations can easily be carried out numerically, using Newton algorithm for instance. Nevertheless, we will make some additional hypothesis, leading to analytical results. In particular, we will suppose that, as in many economic applications, the utility is linear in its argument (here the reliability) so that $U_{s,t}(r) = U_{s,t}+\gamma_{s,t}r$ with $\gamma_{s,t}$ translating the reliability in financial terms, as the monetary value of a reliability unit (so that the utility increases with r) and $U_{s,t}$ being a random variable not depending on r. Let $F^{*}_{s,t}$ be the distribution function of random variable $U_{s,t}$ not depending on r. Then the previous first order conditions become

$$\overline{F}^{*}_{s,t}\left(p_{s,t}-\gamma_{s,t}r\right)+p_{s,t}\frac{\partial\overline{F}^{*}_{s,t}\left(p_{s,t}-\gamma_{s,t}r\right)}{\partial p_{s,t}}=0$$

In particular, if for all s, t $\overline{F}^{*}_{s,t}(p)=\left(\frac{p}{M_{s,t}}\right)^{\alpha_{s,t}+1}$ with $0 \le p \le M_{s,t}$, from the first order conditions we obtain that $p_{s,t}=\gamma_{s,t}r_{s,t}+M_{s,t}\left(\alpha_{s,t}+2\right)^{-1/(\alpha_{s,t}+1)}$ provides the optimal prices and the maximum revenue. In the rest of this paper, we work with this demand function.

3 Extending the Network, based on Requests

The next step is to plan the capacity of the network. The idea is the following: consider a family F of graphs such that $G=(N,E') \in F$, the set N of nodes is the same, but the set of links E' is a subset of possible links E ($E' \subseteq E$).

In order to completely define a network $G=(N,E')$ in our model, we have to choose a technology type for each link $l \in E'$, we express this with the assignment function $a{:}E' \rightarrow T$ (where $a(l)$ means the technology type chosen for the l link, $a(l) \in \mathrm{T}(l)$).

From the network point of view, the goal is to determine the topology $G=(N, E')$ (and the assignment function a) maximizing the benefits $H(G)-\sum_{l\in E'}c_l\left(a(l)\right)$.

Now, we can summarize the formal problem and the notations used throughout the rest of the paper. Inserting in the revenue equation the prices and the demand distribution function we have

$$H(G)=\sum_{(s,t)\in K}\frac{\tilde{\lambda}_{s,t}}{\mu_{s,t}}\left[\gamma_{s,t}r_{s,t}+M_{s,t}\left(\alpha_{s,t}+2\right)^{-1/(\alpha_{s,t}+1)}\right]\frac{\left(\alpha_{s,t}+1\right)}{\left(\alpha_{s,t}+2\right)}.$$

We then have the problem definition

$$\mathit{Maximize}\quad H(G)=\sum_{(s,t)\in K}\frac{\tilde{\lambda}_{s,t}}{\mu_{s,t}}\left[\gamma_{s,t}r_{s,t}+M_{s,t}\left(\alpha_{s,t}+2\right)^{-1/(\alpha_{s,t}+1)}\right]\frac{\left(\alpha_{s,t}+1\right)}{\left(\alpha_{s,t}+2\right)}-\sum_{l\in E'}c_l\left(a(l)\right).$$

4. Genetic algorithm

As the previous formulation does not seem easy to exploit by an exact or numerical algorithm, we use a Genetic Algorithm (GA), which will give an approximate solution. We describe briefly the components of the proposed algorithm:

i. *Encoding*: the genotype (solution encoded) is an array of size given by the number of edges, where we have an allele for each possible link in the network. The alphabet of each allele is an integer between zero and the maximum number of technology types of this link, where zero means that this links does not appear in the solution, and any other value that the link is present ant that we use the corresponding technology type in this link.

ii. *Fitness function*: the objective function defined in Section 3 (the benefits of the network) plus the sum of the maximum costs of all the links (so that fitness is always positive). The computation of the reliabilities (which is an NP-hard problem) is obtained by the generalized antithetic Monte Carlo simulation method proposed by El Khadiri and Rubino (1992).

iii. *Initial population*: generated randomly. Existing links are always included; each non-existing link is selected (or not) according to a Bernoulli variable of parameter 0.8. The type of included links is determined uniformly choosing between possible technology types.

iv. *Stopping criterion*: we tested two criterions, either to fix the number of generations; or to iterate while there is an "improvement" in the solution. The first one was selected, on the basis on tests over calibration problems.

v. *Selection*: "roulette wheel" selection with elitism; the best individual is always included in the next generation, and for the other individuals, the probability of including them in the next generation is proportional to their fitness over the population total fitness.

vi. *Crossover:* single point *Crossover* selecting the crossover point uniformly, and swapping all alleles of the parents between the sampled position and the end of the string. Crossover is applied to two randomly selected strings with a probability p_c (if this does not happen, the parents are copied exactly to the next generation).

vii. *Mutation*: the new value for the current allele is chosen uniformly between zero and the maximum number of technology types of this link. The value zero corresponds to removing the link; the other values include the link with a given technology type.

A remark is that the operations preserve the feasibility of the solutions; this is useful, because feasibility can be hard to maintain in a genetic algorithm when the problem has many constraints.

5 Numerical Illustration

The VTHD (Very High Broadband IP/WDM test platform; see http://www.vthd.org) network is a French project, which main goal is to investigate the applications of a new generation of Internet and Intranet networks. We have used this network for an illustrative application of our method. The VTHD network uses two main technologies types for its links: the backbone part of the network uses a IP/WDM architecture, with STM1/4 and STM16 links (in this work we suppose a probability of failure of 0.01 for these links); the access part of the VTHD network uses Giga-Ethernet links (with a probability of failure of 0.1). Figure 1(a) shows the network for our illustrative example. The same network is shown schematically in Figure 1(b), also representing some feasible additional links (shown as dotted lines).

Specifically, we apply the proposed GA method in three different scenarios. The three problems have the same specification (the same parameters of the demand, utility, etc.), but they differ in the possibilities of network extension. In the first problem (called VTHD1), we evaluate the benefits of extending the backbone of the network with the links shown as dotted lines; in this case the best solution of this problem can be found by enumeration because the network extension has only 32 possibilities. The second and third problems (called VTHD2 and VTHD3 respectively) add in addition the possibility to upgrade the access network with IP/WDM links. The difference between these two problems is the cost of the new possible access links (in VTHD2 problem we use reasonably moderated costs, and in VTHD3 problem we consider worst case high costs). It is very hard to obtain the optimal solution for these two problems, as the solution space is very large (exactly 2^{25} possibilities). The VTHD1 optimal solution is also feasible for these problems, its value providing then a lower bound for their optima.

The experiments were run on a SunFire 280R, with two 1.2 GHz UltraSPARC III Cu processors, 2 GB of main memory, and SolarisTM 8 operating system. The parameters of the algorithm were chosen as follows: mutation rate p_m=0.01, crossover rate p_c=0.95, population size P=100, generation number G=100, generalized antithetic Monte Carlo block size B=100 and number of blocks L =50. These values were chosen on the basis of calibration experiments over a set of ten smaller problems.

The execution times for the three problems are similar. Each mutation takes in average 75.26 milliseconds, and each crossover takes 1.16 seconds. The mutation and the crossover are often executed in the execution of a genetic algorithm (exactly 10000 times the mutation and 5000 times the crossover, because we have a population of size 100, and 100 generations). The selection operator needs the fitness of the population to be computed, therefore, before each selection, we have to calculate the fitness of the new individuals, that implies a reliability estimation. This estimation takes in average 1660.99 milliseconds, the consequence is that the algorithm execution time is approximately 5 hours.

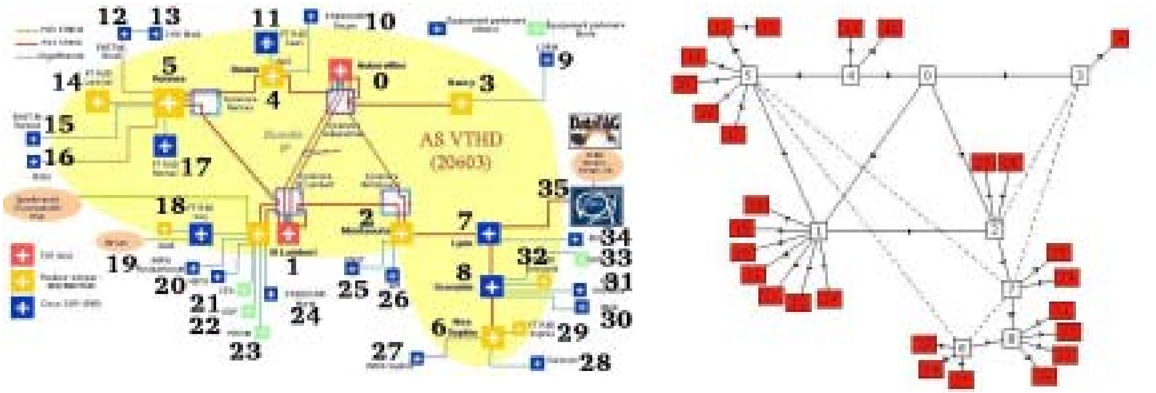

Figure 1: Validation Problem: Very High Broadband IP/WDM test platform.

Table 1 shows the results of the genetic algorithm. We found that for VTHD1 the GA obtains the optimal solution (the values of the benefit are not identical, as their computation by Monte Carlo includes a small statistical error). In the VTHD2 problem, the solution found is quite better than the lower bound, and the selected links are in general quite different. The VTHD3 problem has performance similar to VTHD1; this is good, as the solution found is most probably the optimum.

Problem	Benefit of known feasible solution	Best Fitness of GA	Maximum Cost	Cost	Benefit
VTHD1	1482	1644	160	0	1484
VTHD2	1482	1904	370	130	1534
VTHD3	1482	3743	2260	0	1483

Table 1: VTHD solutions. The maximum cost of the network is the sum, for all links, of the most expensive technology type costs. The Benefit is the objective function.

In Figure 2 we show the evolution of the average fitness and best fitness of the population. For problems VTHD1 and VTHD2, the initial population has already quite good fitness values; that is not the case for VTHD3. All the same, the GA attains quickly good values in the three cases; for VTHD1 and VTHD3, the optimum seems to be found in less than forty generations. In the case of VTHD2, it is difficult to know if the optimum has been attained, but the evolution seems to have stopped after 60 generations.

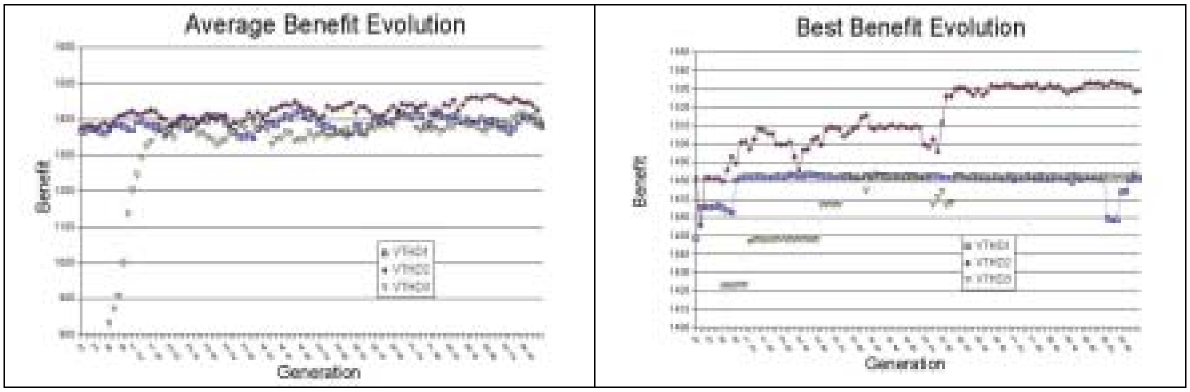

Figure 2: Population average and best benefit for each GA generation.

5 Conclusions and Future Work

In this paper, we have studied a new pricing scheme based on connection reliability, and applied it to fix prices and to extend an already existing network in order to increase the service provider's benefit. To validate our method, we have run the method on three problems inspired by the VTHD (Very High Broadband IP/WDM test platform) network. One of the problems consists in modifying the backbone by incorporating new links. The other two, with a large solution space, additionally upgrade the access network technology of the access links. The genetic algorithm finds the optimal solution of the first problem, a very good solution for one of the extended problems, and the optimal or very near the optimal solution for the other problem.

Note that the fitness evaluation of GA is very hard since it needs a reliability estimation, an NP-hard problem in general. In our work, it is estimated by means of the Generalized Antithetic Monte Carlo simulation method. Our GA evaluates the fitness many times, therefore an important improvement in run time can be attained by diminishing the computing time of each evaluation or the total number of evaluations. As future work, we could try different approaches to solve this problem: a) using efficient upper bounds; this approach is interesting because it can represent Service Level Agreements, based in reliability, in a natural way; b) in the execution of our GA, some reliability estimations might be computed more than once; savings can be done by storing previous computations and avoiding repeating them; c) developing some heuristic methods to estimate the reliability from previous similar estimations (for example, in the literature a random neural network has been proposed).

An important point is the impact on the optimization procedure of the reliability error introduced by the estimation. In experiments not included here for space reasons, we evaluated the standard deviation of the fitness function resulting from reliability estimation in all the execution of our validation problem, and we observed that it is high for some of our instance problems. The computed fitness standard deviation is an upper bound of the uncertainty of the GA method. A refinement of the trade-off between uncertainty in genetic algorithm and reliability estimation is then an important issue for future work, in order to obtain a robust method.

Acknowledgement

This research was supported by the INRIA ARMOR research project, France. The work of Pablo Rodríguez was supported by the French Embassy in Uruguay as part of the French Ministère des Affaires Étrangères scientific cooperation program; and by the "Programa de Jóvenes Investigadores" of CSIC, UDELAR, Uruguay.

References

L.A. DaSilva (2000). Pricing of QoS-Enabled Networks: A Survey. *IEEE Communications Surveys & Tutorials*, 3(2).

M. Falkner, M. Devetsikiotis, and I. Lambadaris (2000). An Overview of Pricing Concepts for Broadband IP Networks. *IEEE Communications Surveys & Tutorials*, 3(2).

T. Henderson, J. Crowcroft, and S. Bhatti (2001). Congestion Pricing. Paying Your Way in Communication Networks. *IEEE Internet Computing*, September/October:85-89.

M. El Khadiri and G. Rubino (1992). A Monte-Carlo method based on antithetic variates for network reliability computations. Technical report, Publication Interne nro 626, IRISA-INRIA, 1992.

B. Tuffin (2003). Charging the Internet without bandwidth reservation: an overview and bibliography of mathematical approaches. *Journal of Information Science and Engineering*, 19(5):765-786.