Call Admission and Routing in Multi-Service Loss Networks

Zbigniew Dziong, *Member, IEEE*, and Lorne G. Mason

Abstract— **A state-dependent policy for call admission and routing in a multi-service circuit-switched network is synthesized. To meet different requirements the objective function is defined as the mean value of reward from the network. Policy iteration is applied to find the optimal control. Assuming link independence the network reward process is decomposed into the set of link reward processes thereby significantly reducing complexity. The approach is implementable for large systems if certain approximations are used. A simulation study shows that the algorithm converges in two iterations, exhibits good traffic efficiency, and provides a flexible tool for performance allocation among services. The approach also constitutes a framework for studying, synthesizing and optimizing other call admission and routing strategies. In particular the results of sensitivity analysis are used to compare the proposed decomposition approach with that developed by F.P.Kelly for optimization of a load sharing policy in telephone networks.**

Keywords— **state-dependent routing, heterogeneous traffic, Markov decision process, decomposition, shadow prices, sensitivity analysis, optimization.**

I. Introduction

Dynamic and adaptive routing, introduced into the North American telephone network in the 1980's, has substantially improved network performance and reliability. Implementation of advanced routing techniques is now in progress worldwide. The next step - extension to integrated-services networks - is the subject of research. These networks are characterized by high link speeds, heterogeneous character of the traffic and performance requirements, and uncertain and variable traffic demand. These features increase the complexity of the traffic management problem and give new prominence to the synthesis of control procedures which ensure efficient operation while maintaining prescribed service levels for the different traffic classes.

One promising approach to this problem involves the synthesis of adaptive control algorithms for call set-up, (call admission and routing), which respond to measured network conditions. A number of approaches to the adaptive routing of calls in telephone networks have been proposed in the literature. These methods range from decentralized adaptive schemes employing learning automata [1], through centralized time-variable schemes [2] to adaptive routing procedures based on the least loaded path (LLP) approach [3, 4, 5, 6]. A framework for an application of the strategies described in [2, 5] to multiservice integrated networks is described in [7]. In this scheme each service has access to dedicated and shared bandwidth so the service performance can be controlled. There is also a substantial literature on call admission control in one link multiservice networks. These works vary from some optimal models [8, 9, 10] to simplified schemes based on threshold type policies [11, 12, 13].

More recently two new elements have been added to the algorithms for call admission and routing in telephone networks: Markov decision theory and revenue maximization. In [14] and [15], Markov decision theory was employed to compute a state dependent routing policy off-line by executing a single policy iteration step. In [16] it was shown that a modified version of this scheme is more efficient than the least loaded path approach. Kelly [17] has introduced the notion of link shadow prices to the problem of decentralized adaptive load sharing with revenue maximization objective (a simplified version of this concept was also presented in [18]). The link shadow price can be interpreted as an average price paid for carrying a call on the link. In [19], it was shown how this concept can be generalized to state-dependent link shadow prices. In this case the performance was also superior to the least loaded path approach. Moreover it was shown that the call revenue parameters provide a means for controlling almost independently and continuously the grade of service (GOS) of different traffic classes. Note that this feature is crucial in the control of multi-service circuit-switched networks where call classes with different bandwidth requirements can encounter very different grades of service.

The new advantages, achieved by the application of revenue maximization and Markov decision theory to telephone network control, motivated us to extend and generalize these concepts to multi-service circuit-switched loss networks. In the paper we formulate the call admission and routing problem, for such networks, as maximization of the reward from all carried calls (each call is characterized by a reward parameter). In general this problem can be solved by applying the policy iteration algorithm from Markov decision theory (Section II). However, the exact state model leads to enormous complexity, putting computation of the control beyond reach for practical cases. To achieve an implementable solution we propose (Section III) a novel decomposition of the Markov decision problem. Namely it is shown that, under the statistical link independence assumption, the network reward process can be de-

Paper approved by Jeffrey M.Jaffe, the Editor for Routing and Switching of the IEEE Communications Society. Manuscript received: October 28, 1991; revised May 1, 1992; September 16, 1992. This paper was presented at the Proc. of the 28th IEEE Conference on Decision and Control, Tampa, FL, December 1989 and the Proc. of IEEE INFOCOM'92 The Conference on Computer Communications, Florence, Italy, May 1992.

The authors are with INRS-Telecommunications, Verdun, P.Q., Canada H3E 1H6.

IEEE Log Number 9401603.

composed into a set of separable link reward processes. In this model the control decision is based on state-dependent link shadow prices interpreted as a predicted price for seizing link bandwidth by the call. Despite decomposition, the control policy is still evaluated by the complete policy iteration algorithm. This is in contrast to the model presented in [14, 15] where only one iteration can be executed. The important element of the proposed scheme is that the analytical part of the model is fed by real time traffic measurements of current flow distribution in the network. This feature also implies that the network control will track or adapt to time variable traffic demand.

Several link models for the evaluation of state-dependent shadow prices are described in Section IV. They vary from the exact solution based on the value iteration procedure to a simplified model using the *"near complete decomposability"* feature and recurrence solution. To evaluate link arrival rates a simplified model using the traffic measurements is applied.

In Section V it is shown that the call average of link shadow price can be used to evaluate sensitivity of the average reward from the network with respect to arrival rates. This result is later used to reduce the error in traffic flow distribution (caused by the independence assumption) and speed up the convergence in the control model. Sensitivity analysis results are also used to show the relationship of the proposed approach, when reduced to the load sharing strategy in telephone networks, to the work of Kelly [17]. It is shown that while in the proposed approach the network reward process is decomposed into separable link processes, in Kelly's approach only the network Markov process is decomposed.

In the final part of the paper (Section VI) the important characteristics of the algorithm are studied. The results confirm the predicted ability to provide almost continuous and independent control of each call class GOS by simply changing the call reward parameters. To test the efficiency of the proposed algorithm we compare it with a generalized version of the least loaded path strategy, known to be very efficient for telephone networks. In all cases tested the proposed algorithm provides better utilization of the network resources. We indicate inherent reasons justifying this result.

In the conclusions (Section VII) we underline new contributions of the paper - the main being synthesis of a control model for multiservice networks where all basic functions, including control of each service GOS, are included and optimized in one homogeneous, theoretically based model fed by real time measurements. Also some further areas of investigation are indicated.

Although the presentation focuses on multi-service networks most results and conclusions are of course applicable to the special case of one service networks (e.g. telephone networks).

II. Problem Formulation

We describe the network as a set of nodes and a set of trunk groups connecting the nodes. The network is offered many classes of calls. The j-th class is characterized by the following: origin-destination (OD) node pair, number of required channels, d_j, intensity of arrival process (assumed to be Poissonian), λ_j, mean holding time (assumed to be exponentially distributed), μ_j^{-1}, set of alternative paths (in general this set can contain all feasible paths), W_j, and reward parameter $r_j \in (0, \infty)$ which can be interpreted as the average reward for carrying the j-th class call (the reward rate from a carried call is given by $q_j = r_j \mu_j$). The network operates in a lost call mode and when a call is accepted we assume instantaneous call set up.

The problem addressed in the paper can be formulated as follows: find the optimal routing policy π^* which maximizes the mean value of reward from the network defined as:

$$\overline{R}(\pi) = \sum_j r_j \overline{\lambda}_j \tag{1}$$

where $\overline{\lambda}_j$ denotes the average rate of accepted j-th class calls (the process is assumed to be stationary). In general this problem can be solved within the framework of the theory of continuous-time Markov Decision Processes (MDP). In the following we present this solution.

The state of the considered system can be described by a matrix $\mathbf{z} = \{z_j^k\}$ where z_j^k denotes the number of j-th class calls carried on the k-th path from the W_j and $\mathbf{z} \in Z$. For each state the rate of reward from the system, $q(\mathbf{z})$, is given by

$$q(\mathbf{z}) = \sum_j \sum_k r_j \, z_j^k \, \mu_j. \tag{2}$$

In the case of call arrival the state transition is described as $\mathbf{z} \rightarrow \mathbf{z} + \mathbf{\Delta}_j(\mathbf{z}, \pi)$ where $\mathbf{z} + \mathbf{\Delta}_j(\mathbf{z}, \pi)$ denotes the state after accepting the j-th type call on path k , recommended by policy π, in state $\mathbf{z}$. In cases where either all paths from W_j are blocked or it is not efficient to carry the call, the decisions are defined by $\mathbf{\Delta}_j(\mathbf{z}, \pi) = \{0\}$. In the case of call departure the state transition is described as $\mathbf{z} \rightarrow \mathbf{z} - \delta_j^k$ where $\mathbf{z} - \delta_j^k$ denotes state after the departure of the j-th type call from path k in state $\mathbf{z}$. The rates of the transitions are λ_j and $z_j^k \cdot \mu_j$, respectively.

From the MDP theory [20, 21] it follows that since our system is ergodic, the optimal policy π^* is deterministic and can be found by applying one of the well known algorithms (eg. policy iteration, linear programming, value iteration). We have chosen the policy iteration algorithm resulting in the following iteration cycle:

- *For given policy π, solve the set of value-determination equations (for relative values $v(\mathbf{z}, \pi)$; $\mathbf{z} \in Z$)*

$$\begin{aligned} \overline{R}(\pi) \;\; = \;\; & q(\mathbf{z}) + \sum_j \lambda_j \left[v(\mathbf{z} + \mathbf{\Delta}_j(\mathbf{z}, \pi), \pi) - v(\mathbf{z}, \pi) \right] \\ & + \sum_j \sum_k z_j^k \, \mu_j \left[v(\mathbf{z} - \delta_j^k, \pi) - v(\mathbf{z}, \pi) \right] \end{aligned} \tag{3}$$

by setting the relative value for an arbitrary reference state $\mathbf{z}_r$ to zero.

- *For each state $\mathbf{z}$ find the alternative set of decisions,*

$\boldsymbol{\Delta}_j(\mathbf{z},\pi')$), *that maximizes the expression*

$$q(\mathbf{z}) + \sum_j \lambda_j \left[v(\mathbf{z}+\boldsymbol{\Delta}_j(\mathbf{z},\pi'),\pi) - v(\mathbf{z},\pi)\right] + \sum_j \sum_k z_j^k \mu_j \left[v(\mathbf{z}-\delta_j^k,\pi) - v(\mathbf{z},\pi)\right] \tag{4}$$

using the relative values from the previous policy. This set of decisions constitute an improved policy π' to be used again in the first step.

The theory of MDP ensures that starting from an arbitrary initial policy this procedure converges to π^* in a finite number of iterations. However, for most telecommunication networks, the policy iteration procedure based on the exact network state description is intractable due to the enormous cardinality of the state and policy spaces. In the following section we propose an approach where the network reward process is decomposed into a set of separable link reward processes. While this approximation reduces computational and memory requirements to manageable levels, the policy iteration algorithm is preserved.

For further consideration it is convenient to introduce the notion of the path net-gain, $g_j(\mathbf{z},\pi)$, defined as

$$g_j(\mathbf{z},\pi) = v(\mathbf{z}+\boldsymbol{\Delta}_j(\mathbf{z},\pi),\pi) - v(\mathbf{z},\pi). \tag{5}$$

It can be shown that the path net-gain can be also expressed as

$$g_j(\mathbf{z},\pi) = \lim_{T\to\infty}\left[R(\mathbf{z}+\boldsymbol{\Delta}_j(\mathbf{z},\pi),\pi,T) - R(\mathbf{z},\pi,T)\right] \tag{6}$$

where $R(\mathbf{z},\pi,T)$ denotes the expected reward from the network in the interval (t_0, t_0+T), assuming state $\mathbf{z}$ in t_0. Note that the separable form of the objective function (4) to be maximized in the second step of the iteration procedure, assures that this step is equivalent to the separate maximization of $g_j(\mathbf{z},\pi)$ over π, for each state $\mathbf{z}$ and call class j, pair.

III. Decomposition of MDP

First we assume that link arrivals are state dependent Poisson streams and that link state distributions are statistically independent. These assumptions are commonly made in network performance analysis. In particular, they imply that a call connected on a path consisting of l links is decomposed into l independent link calls characterized by the same mean holding time as the original call. Then the Markov process for a given policy π can be described separately for each link in terms of the link state $\mathbf{x} = \{x_j\}$ and the transition rates defined by the link arrival rates $\lambda_j^s(\mathbf{x},\pi)$ and departure rates μ_j, where x_j denotes the number of j-th class calls carried on the considered link (to simplify notation, we assumed that alternative paths for j-th class calls have no common links - this is not a limitation of the approach). Concerning the evaluation of $\lambda_j^s(\mathbf{x},\pi)$, it can be done in two ways. One is to develop an analytical model for performance analysis of the network with the given routing policy π. The second possibility is to estimate $\lambda_j^s(\mathbf{x},\pi)$ based on some simple statistics measured in the network. We have chosen the second option due to its smaller complexity and automatic adaptation to changes in the traffic patterns. The details of this approach are given in Section IV.A.

Although the above mentioned assumptions provide a decomposition of the Markov process, it is not sufficient to decompose the analysis of the Markov decision problem. To do that we propose decomposition of the network reward process into a set of separable link reward processes. This can be done by dividing the reward parameter of a call offered to a multilink path among the link calls, so that each link call is characterized by the link call reward parameter, $r_j^s(\pi)$. It is clear that the division rule should provide maximization of the mean value of reward from the network with the obvious constraint:

$$r_j = \sum_{s\in k} r_j^s(\pi) \tag{7}$$

where k denotes the path chosen by policy π for carrying the j-th type call. The division of the reward parameter is discussed in Section IV.A.

Now, for a given routing policy, each link reward process can be described independently by the set $\{\lambda_j^s(\mathbf{x},\pi), r_j^s(\pi), \mu_j\}$ enabling definition of the link net-gain, $g_j^s(\mathbf{x},\pi)$, as the expected increase in the reward from the link caused by accepting the j-th class link call:

$$g_j^s(\mathbf{x},\pi) = \lim_{T\to\infty}\left[R^s(\mathbf{x}+\delta_j,\pi,T) - R^s(\mathbf{x},\pi,T)\right] \tag{8}$$

where $R^s(\mathbf{x},\pi,T)$ denotes the expected reward from the link in the interval (t_0, t_0+T), assuming state $\mathbf{x}$ in t_0, and δ_j is J-vector with 1 in the position j and zeros in all other positions. Observe that under the link independence assumption we have

$$\lim_{T\to\infty} R(\mathbf{y},\pi,T) = \sum_s \lim_{T\to\infty} R^s(\mathbf{x},\pi,T) \tag{9}$$

where $\mathbf{y} = \{\mathbf{x}\}$ denotes the network state in the decomposed model. Thus from (6, 8) it follows that the path net-gain for the decomposed model is given by

$$g_j(\mathbf{y},\pi) = \sum_{s\in k} g_j^s(\mathbf{x},\pi). \tag{10}$$

This separable form of the path net-gain constitutes our basis for decomposition of the Markov decision problem since it was shown in the previous section that the values of path net-gain define the policy improvement procedure.

For subsequent simplifications of our model it is convenient to define a state-dependent link shadow price $p_j^s(\mathbf{x},\pi)$ which is related to the link net gain and the link call reward by the equation

$$p_j^s(\mathbf{x},\pi) = r_j^s(\pi) - g_j^s(\mathbf{x},\pi). \tag{11}$$

This value can be interpreted as the expected price for accepting a j-th type call on link s in state $\mathbf{x}$. Notice that from (7, 10, 11) it follows that the path net-gain from carrying the j-th type call on the path k can be expressed as

$$g_j(\mathbf{y},\pi) = r_j - \sum_{s\in k} p_j^s(\mathbf{x},\pi). \tag{12}$$

The models for evaluation of the link shadow prices (and link net-gains) are given in Section IV.B.

Using the concept of state dependent link shadow prices, the policy iteration procedure from the previous section can be rewritten in the following steps:

- *Collect the statistics in the network operating under a given routing policy π, then evaluate $\lambda_j^s(\mathbf{x},\pi)$ and, using these values, compute the new values of the link shadow prices $p_j^s(\mathbf{x},\pi)$,*
- *For each arrival of the j-th class call find the maximum net-gain over all feasible paths*

$$g_{max} = \max_{k\in W_j}[r_j - \sum_{s\in k} p_j^s(\mathbf{x},\pi)] \tag{13}$$

using the new values of link shadow prices, if g_{max} is positive, carry the call on the path giving the maximum net-gain, otherwise reject the call.

Assuming that starting from an arbitrary policy this procedure (henceforth called MDPD) converges to a limit policy π^*, we could treat this solution as the optimal policy if all used assumptions are exact. Since there are approximations, policy π^* is in general suboptimal.

The main advantage of the presented decomposition approach (compared to the exact model) is that the evaluation of the path net-gain is decomposed into link analysis problems (state space reduction) and that the policy need not to be stored for all network states but instead the decisions can be easily computed at the instant of call arrivals on the basis of link shadow prices (memory requirement reduction). It is also important that the evaluation of shadow prices is based on the real-time measurement which implies that the routing policy will track or adapt to a time variable traffic demand.

In the conclusion of this section we indicate the methodological difference between the proposed approach and the models for control of telephone networks, with the objective to maximize total traffic, presented in [14, 15]. Namely in [14, 15] the separability of link shadow prices (link costs in [14, 15]) is achieved by applying direct routing as the initial policy in the standard policy iteration algorithm. Thus by neglecting multilink flows the link costs can be easily evaluated at the price of restricting the policy iteration procedure to one iteration since the link distributions cannot be changed. This idea could be applied to the multi-service problem as well (in this case $\lambda_j^s(\mathbf{x},\pi) = \lambda_j$ and $r_j^s(\pi) = r_j$). Nevertheless as it will be shown in Section VI the performance of this option can be significantly worse than that of the MDPD strategy.

IV. Link Shadow Price Evaluation

A. Link Call Parameters

The approach assumes that the following statistics are measured in the network (for each call class): average arrival rate of offered calls, $\lambda_j(\pi)$, and average rate of calls accepted on each alternative path, $\overline{\lambda}_j^k(\pi)$. Observe that these statistics can be also very helpful in network management and planning.

The evaluation of the state dependent link arrival rates is based on the assumption that the arrival rates seen by links (in the non blocking states) under a state-dependent routing can be approximated by a load sharing model where the arrival rate of stream offered to the k-th path, $\lambda_j^k(\pi)$, is Poissonian (similar assumptions were made in the performance models for networks with state-dependent routing presented in [24, 25]). We assume that $\lambda_j^k(\pi)$ is proportional to the rate of accepted j-th type calls on path k:

$$\lambda_j^k(\pi) = \lambda_j \frac{\overline{\lambda}_j^k(\pi)}{\sum_{k\in W_j}\overline{\lambda}_j^k(\pi)}. \tag{14}$$

Then under the link independence assumption the link state dependent arrival rates in the non blocking states can be found from the path model and are given by

$$\lambda_j^s(\mathbf{x},\pi) = \lambda_j^k(\pi)\cdot f_j^s(\mathbf{x},\pi) \prod_{c\in k\setminus\{s\}} (1 - B_j^c(\pi)) \tag{15}$$

where $B_j^c(\pi)$ denotes the probability that link c has not enough free capacity to accept the j-th type call (blocking state) and $f_j^s(\mathbf{x},\pi)$ denotes a filtering probability defined as

$$f_j^s(\mathbf{x},\pi) = P\{\sum_{c\in k\setminus\{s\}} p_j^c(\mathbf{x},\pi) < r_j - p_j^s(\mathbf{x},\pi)|\overline{B}_j\} \tag{16}$$

where $\overline{B}_j$ denotes condition that neither link is in the blocking state (note that $p_j^s(\mathbf{x},\pi)$ is constant in (16)). In other words $f_j^s(\mathbf{x},\pi)$ is the probability that the path net-gain is positive (on condition that there is enough path capacity to carry the call). This probability can be computed using the link state distributions. Note that for evaluation of the link shadow prices we do not need to define the link arrival rates for the link blocking states. Nevertheless to simplify presentation of the link model we assume $\lambda_j^s(\mathbf{x},\pi) = 0$ in these states.

Concerning evaluation of the link call reward parameters, $r_j^s(\pi)$, the rule for allocation of r_j among the path's links can have an influence on the average reward from the network. The exact solution maximizing the reward from the network is quite complex since the optimization procedure would require a network performance model. Nevertheless the investigation presented in [26] based on an analytical path model indicated that in the practical range of parameters the reward from the network is quite insensitive to the division rule. That is why we focus on a simple solution which, due to its economical interpretation, can be very attractive in network management and dimensioning.

TABLE I
REWARD LOSSES [%] FOR DIFFERENT REWARD ALLOCATION RULES

	D1	D2	D3
nominal conditions			
W7N	0.96 ± .15	1.03 ± .12	1.02 ± .17
N7N	1.44 ± .09	1.42 ± .09	1.41 ± .09
N11N	0.51 ± .07	0.51 ± .06	0.53 ± .06
overload conditions			
W7N	4.84 ± .32	4.83 ± .30	4.73 ± .32
N7N	6.12 ± .10	6.10 ± .10	6.15 ± .10
N11N	5.21 ± .10	5.19 ± .11	5.23 ± .11

Namely we assume that the reward parameter $r_j^s(\pi)$ assigned to the s-th link should be proportional to the call average link shadow price, $\overline{p}_j^s(\pi)$, paid for carrying the j-th type link call on this link:

$$r_j^s(\pi) = r_j \frac{\overline{p}_j^s(\pi)}{\sum_{c\in k} \overline{p}_j^c(\pi)} \tag{17}$$

where $\overline{p}_j^s(\pi)$ is defined by

$$\overline{p}_j^s(\pi) = E_c[p_j^s(\mathbf{x},\pi)] = \sum_{\mathbf{x}\in X} Q_j(\mathbf{x}) p_j^s(\mathbf{x},\pi) \tag{18}$$

and $Q_j(\mathbf{x})$ denotes the probability that the j-th type call is accepted in state $\mathbf{x}$. Note that $\overline{p}_j^s(\pi)$ can be estimated in a real network by averaging the values of link shadow price at the instants of the j-th type calls arrivals. It can be shown that the proposed rule is optimal in the case of a fully symmetrical path and in the case where $\overline{p}_j^s(\pi) = 0$.

To illustrate the conclusion that in the practical range of parameters the reward from the network is not sensitive to the division rule we compare performance (expressed as the network reward losses, $H = 1 - \overline{R}/(\sum_j r_j \lambda_j)$), in three network examples (described in Section VI) for three different division rules. The first rule (D1) is described above. The second (D2) assumes that all link call reward parameters are equal to each other. The last one (D3) is an adaptation of the reward allocation rule from the model developed by Kelly [17] for the load sharing strategy:

$$r_j^s(\pi) = r_j - \sum_{c\in k\setminus\{s\}} \overline{p}_i^c(\pi) \tag{19}$$

In this case the sum of link call reward parameters is not equal to the call reward parameter (equation (7) does not hold). The explanation of this fact and relation of the MDPD approach to Kelly's model are described in Section V. The results from simulation model are presented in Table I. The performance of all three versions is very close to each other and falls within the confidence interval of the other versions. This result suggests that from a control performance point of view the simple rule D2 is sufficient in the practical range of parameters. Nevertheless as indicated in [23] the natural economical interpretation of the D1 rule can be very attractive from a management, planning and dimensioning point of view. It can be also shown that in some limiting cases the performance with the D2 rule is worse than the one with the D1 rule.

B. Exact Link Models

Once the values of link arrival rates, $\lambda_j^s(\mathbf{x},\pi)$, and link call reward parameters, $r_j^s(\pi)$, are given, the link net-gains can be evaluated by solving the following set of equations (achieved by applying (3) and (5) to the link reward process)

$$\begin{aligned}\overline{R}^s(\pi) &= q(\mathbf{x}) + \sum_j \lambda_j^s(\mathbf{x},\pi)\, g_j^s(\mathbf{x},\pi) \\ &\quad - \sum_j x_j\, \mu_j\, g_j^s(\mathbf{x}-\delta_j,\pi) \quad ; \mathbf{x}\in X^s \end{aligned} \tag{20}$$

where $\overline{R}^s(\pi)$ denotes average reward from the link and $q(\mathbf{x}) = \sum_j r_j^s(\pi) x_j \mu_j$ is the rate of link reward in state $\mathbf{x}$.

An attractive alternative is the value-iteration algorithm [21] which in general is a convenient method for solving large Markov problems due to numerical simplicity (see e.g. [22]). Since this algorithm is applicable directly only for discrete time Markov processes we must first apply the uniformization technique with certain average length of the transition time, τ [21]. Then the basic recurrence for our link model is stated as follows:

$$\begin{aligned} V_n^s(\mathbf{x},\pi) &= q(\mathbf{x})\cdot\tau \\ &\quad + \sum_j \lambda_j^s(\mathbf{x},\pi)\,\tau\,[V_{n-1}^s(\mathbf{x}+\delta_j,\pi) - V_{n-1}^s(\mathbf{x},\pi)] \\ &\quad + \sum_j x_j\,\mu_j\,\tau\,[V_{n-1}^s(\mathbf{x}-\delta_j,\pi) - V_{n-1}^s(\mathbf{x},\pi)] \\ &\quad + V_{n-1}^s(\mathbf{x},\pi) \quad ; \mathbf{x}\in X^s \end{aligned} \tag{21}$$

where the value function, $V_n^s(\mathbf{x},\pi)$, can be interpreted as the expected reward from the link within n transition periods assuming state $\mathbf{x}$ at the beginning of the considered time and terminal reward of $V_0^s(\mathbf{x},\pi)$ at the end of this time. It can be proved [21] that starting from an arbitrarily chosen $\{V_0^s(\mathbf{x},\pi)\ ; \mathbf{x}\in X\}$ the difference $V_n^s(\mathbf{x},\pi) - V_{n-1}^s(\mathbf{x},\pi)$ will be as close as needed to the average reward from the link, $\overline{R}^s(\pi)\cdot\tau$, for sufficiently large n. It can be shown that having the value functions, the link net-gain can be expressed as

$$g_j^s(\mathbf{x},\pi) = \lim_{n\to\infty} [V_n^s(\mathbf{x}+\delta_j,\pi) - V_n^s(\mathbf{x},\pi)] \tag{22}$$

It is important that in the particular case of one call class the set of equations [20] can be rewritten in the form of two recurrence relations which provide a very efficient solution for link net-gains. Namely after solving recurrences

$$u(x) = \frac{1 + x\cdot\mu\cdot u(x-1)}{\lambda(x,\pi)} \quad ; x = 1,..,N-1 \tag{23}$$

$$w(x) = \frac{x\cdot\mu\cdot w(x-1) - q(x)}{\lambda(x,\pi)} \; ; x = 1,..,N-1 \tag{24}$$

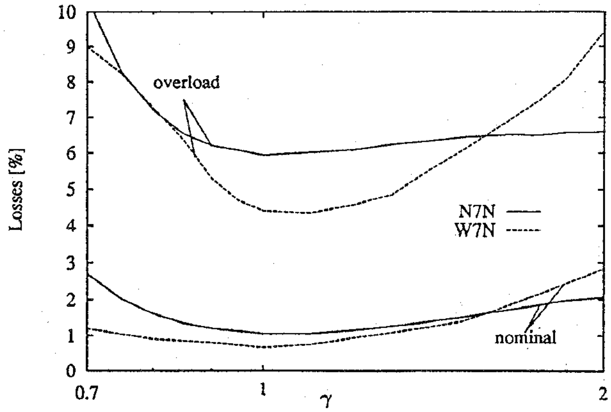

Fig. 1. Reward losses vs. shadow price factor, γ.

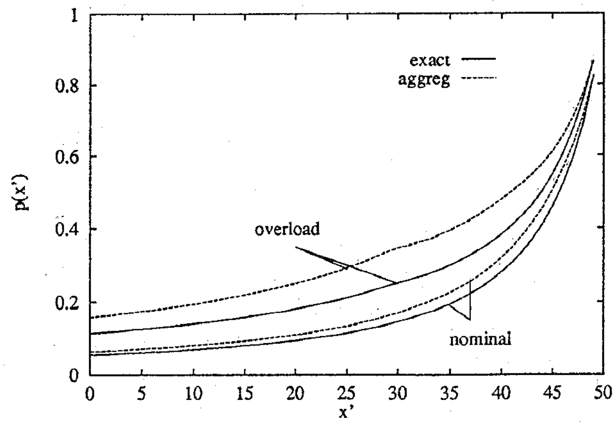

Fig. 2. Shadow price vs. link state.

with initial values $u(0) = 1/\lambda(0, \pi)$ and $w(0) = 0$, we have

$$\overline{R}^s(\pi) = \frac{q(N) - N \cdot \mu \cdot w(N-1)}{1 + N \cdot \mu \cdot u(N-1)} \tag{25}$$

$$g^s(x, \pi) = \overline{R}^s(\pi) u(x) + w(x) \quad ; x = 0, .., N-1 \tag{26}$$

where N denotes link capacity. This solution can be viewed as a generalization of the models for link cost evaluations given in [14, 15], since it covers the cases of state-dependent arrival rates. The same feature is possessed by another recurrence solution given in [19] which was derived analogously to the equations for the mean of the first passage time of the Markov chain.

C. Link Model Simplifications

Despite significant reduction of state space in the MDPD model, one can still encounter some numerical problems or time constraints during the solution of (20) or (21) if the link state space X is very large. One obvious remedy is to simplify the link model. Such an approach is encouraged by the fact that the performance of the control is hardly affected even if the optimal values of shadow prices are changed by several percent. This feature is illustrated in Fig.1 where we presented the performance of the networks W7N and N7N (described in Section VI) versus parameter γ by which all shadow prices in the decision algorithm are multiplied. The function $H(\gamma)$ is shallow in a quite large surrounding of the optimal points. In the following we present several simplifications that can significantly reduce the cardinality of the link state space.

1) Aggregation of Link Call Classes: Let us construct a modified link reward process in which the link call classes with the same bandwidth requirement and mean holding time are aggregated into one class i with an average reward parameter defined as

$$r_i^s(\pi) = \frac{\sum_{j \in i} r_j^s(\pi) \overline{\lambda}_j^s(\pi)}{\sum_{j \in i} \overline{\lambda}_j^s(\pi)} \tag{27}$$

where $\overline{\lambda}_j^s(\pi)$ denotes the average rate of j-th class calls accepted on the link. In general the aggregated reward process should be statistically close to the original process so one can expect that

$$p_j^s(\mathbf{x}, \pi) \cong p_i^{s'}(\mathbf{x}, \pi) \quad ; \; j \in i \tag{28}$$

In fact it can be shown (for details see [27]) that in many cases this relation becomes equality. In particular this is true when all streams have Poissonian not state-dependent arrival rates. The results for an opposite case are presented in Fig.2 and Table II (the studied network example N7S is described in Section VI). The control version with priority for direct link is applied (for details see Section VI) so each direct link stream has steady Poissonian arrival rate and multilink streams are state-dependent (modeled according to (15)). The values of link shadow prices $p(x')$, where x' is the total number of carried calls, presented in Fig.2 indicate that the model based on aggregation overestimates the shadow price and the error (compared to the exact state description) is in the range of few percent of the average link call reward (in this case the shadow prices are the same for direct and multilink calls). Despite this overestimation the differences in the reward losses H, presented in Table II, are within the confidence intervals (simulation model).

TABLE II
REWARD LOSSES, H, AND SHADOW PRICE ERROR, $\overline{\epsilon}$, FOR DIFFERENT LINK MODELS

link model	λ_d	λ_m	$\overline{\epsilon}$	H [%]
		N7S - nominal conditions		
exact	42.0	3.6-0.1		0.28 ±.05
aggreg.	42.0	3.6-0.1	0.036	0.30 ±.05
SAR	42.0	3.6	0.001	0.27 ±.03
		N7S - overload conditions		
exact	46.2	6.6-0.0		4.37 ±.12
aggreg.	46.2	6.6-0.0	0.063	4.43 ±.16
SAR	46.2	6.6	0.053	4.38 ±.12

- λ_d, λ_m denote total arrival rates of direct and multilink streams offered to a link, respectively.

Additionally the call average of the absolute error in the values of shadow prices, $\overline{\epsilon} = E_c[p_i^{s'}(\mathbf{x}, \pi^*) - p_j^s(\mathbf{x}, \pi^*)]$, is given (based on link analytical model).

2) Steady Arrival Rates: In the model for link arrival rates evaluation, (15), the largest part of the complexity is attributed to the evaluation of filtering probabilities. To assess the importance of these probabilities from the network performance point of view we investigated a simplified model (henceforth called SAR) where the filtering probability is assumed to be equal to one for all states. The results for the network example N7S are presented in Table II. The comparison between the exact and SAR link models shows that although the error in shadow price values, $\overline{\epsilon}$, is as large as a few percent of the call reward parameter value ($r_j = 1$), its influence on the network reward losses, H, is negligible.

3) Decomposition of the link model: Let us divide the link call classes into two groups: narrow-band (NB) class with $d_n = 1$ and wide-band (WB) classes with $d_i > 1$. The main idea of the link model decomposition is based on the assumption that since the transition rates of wide band calls are expected to be significantly smaller than the ones for NB calls, the NB calls reach the steady state distribution instantaneously for each state of WB calls. This feature known also as *near complete decomposability* allows to analyse the stationary properties of the system separately for NB Markov process (for each state of WB calls) and WB Markov process. After applying this decomposition technique, the shadow prices for the NB calls can be evaluated from the recurrence solution (23-26) and the evaluation of the shadow prices for WB calls is simplified by neglecting the NB calls (in the case of one WB call class also the recurrence solution can be used). The mutual influence of the two processes is taken into account by reducing the arrival rates of WB calls (blocking caused by NB calls), by adding the average reward from NB calls to the rate of reward from WB calls and by applying the dynamic bandwidth allocation (DBA) [13] to protect WB calls against NB calls. More details on this model can be found in [27, 13]. The efficiency of this approach (henceforth called DLM) is illustrated in Section VI.

V. Average Shadow Price as a Sensitivity Measure

Let us define the average path net-gain, $\overline{g}_j(\pi)$, in the exact model as

$$\overline{g}_j(\pi) = E_{\mathbf{z}}[g_j(\mathbf{z}, \pi)] = \sum_{\mathbf{z} \in Z} Q(\mathbf{z}, \pi) g_j(\mathbf{z}, \pi) \qquad (29)$$

where $Q(\mathbf{z}, \pi)$ denotes the state probability and $g_j(\mathbf{z}, \pi) = 0$ for $\mathbf{z}$ where call j is rejected.

Theorem 1 : *In the exact network model the derivative of average network reward with respect to arrival rate of class j is given by*

$$\frac{d\overline{R}(\pi)}{d\lambda_j} = E_{\mathbf{z}}[g_j(\mathbf{z}, \pi)]. \qquad (30)$$

Proof : From the sensitivity analysis models presented in [28, 29], it follows that in the case of our system the right-hand derivative is given by

$$\frac{d^+\overline{R}(\pi)}{d\lambda_j} = \lim_{T \to \infty} E[\int_{t_0-T}^{t_0+T} (q'(\mathbf{z}, \pi) - q(\mathbf{z}, \pi))dt] \qquad (31)$$

where $q'(\mathbf{z}, \pi)$ denotes the reward rate of the process with one additional call added randomly at time t_0. By using (6) it can be shown that (31) is equivalent to

$$\frac{d^+\overline{R}(\pi)}{d\lambda_j} = E_{\mathbf{z}}[g_j(\mathbf{z}, \pi)] \qquad (32)$$

Analogous proof holds for the left-hand derivative. In this case one call is removed randomly from the system. ■

For further considerations it is convenient to define the derivative of the average network reward with respect to the average rate of calls accepted on the path:

$$\frac{d\overline{R}}{d\overline{\lambda}_j^k} = \lim_{\Delta\lambda_j \to 0} \frac{\overline{R}(\overline{\lambda}_j^k + \Delta\overline{\lambda}_j^k) - \overline{R}(\overline{\lambda}_j^k)}{\Delta\overline{\lambda}_j^k} \qquad (33)$$

where $\Delta\overline{\lambda}_j^k$ corresponds to the change of the average rate of calls accepted on path k in case λ_j is changed by $\Delta\lambda_j$ but all calls from $\Delta\lambda_j$ are ignored except ones that would be carried on path k. In the same manner as proof for Theorem 1 one can show that

$$\frac{d\overline{R}}{d\overline{\lambda}_j^k} = E_c[g_j^k(\mathbf{z}, \pi)] = \sum_{\mathbf{z} \in Z} Q_j^k(\mathbf{z}, \pi) g_j^k(\mathbf{z}, \pi) \qquad (34)$$

where $Q_j^k(\mathbf{z}, \pi)$ denotes the probability that the j-th type call offered to the k-th path is accepted in state $\mathbf{z}$.

It can be shown that in the MDPD model we have

$$\overline{g}_j(\pi) = E_{\mathbf{y}}[g_j(\mathbf{y}, \pi)] = \sum_{k \in W_j} \frac{\overline{\lambda}_j^k}{\lambda_j}(r_j - \sum_{s \in k} \overline{p}_j^s(\pi)) \qquad (35)$$

Since the Theorem 1 is valid also for the MDPD model, based on (30, 34, 35) we have

$$\frac{d\overline{R}(\pi)}{d\lambda_j} = E_{\mathbf{y}}[g_j(\mathbf{y}, \pi)] = \sum_{k \in W_j} \frac{\overline{\lambda}_j^k}{\lambda_j}(r_j - \sum_{s \in k} \overline{p}_j^s(\pi)) \qquad (36)$$

and

$$\frac{d\overline{R}}{d\overline{\lambda}_j^k} = E_c[g_j^k(\mathbf{y}, \pi)] = r_j - \sum_{s \in k} \overline{p}_j^s(\pi) \qquad (37)$$

The results from the sensitivity analysis can be applied to optimization of the load sharing policy, π^f, where the j-th type call is offered to the k-th path with probability h_j^k so $\lambda_j^k = h_j^k \lambda_j$. We do not impose any policy concerning call admission to the path. In case the call cannot be accepted by the chosen path it is lost. The reason for considering this scheme is that conclusions from its analysis can be helpful in analyzing and improving other schemes.

Theorem 2 : *Equalization of the derivatives* $\frac{\partial \overline{R}(\pi^f)}{\partial \lambda_j^k}$ *on all paths used by call class j, is a necessary condition to maximize* $\overline{R}(\pi^f)$ *over the load sharing probabilities,* $\{h_j^k\}$.

Proof : By changing the optimization variables from h_j^k to λ_j^k and transforming the first order Kuhn-Tucker conditions one can achieve the following optimality conditions

$$\frac{\partial \overline{R}(\pi^f)}{\partial \lambda_j^k} = -v_j \quad ; \quad k \in \{k|\lambda_j^k > 0\} \tag{38}$$

where v_j denotes the associated lagrangian multiplier. This condition also implies that a path with negative $\frac{\partial \overline{R}(\pi^f)}{\partial \lambda_j^k}$ should not be offered any traffic of type j. ■

Note that due to (36), in the MDPD model, the optimality conditions can be rewritten as follows

$$\frac{\partial \overline{R}(\pi^f)}{\partial \lambda_j^k} = \frac{\overline{\lambda}_j^k}{\lambda_j^k}(r_j - \sum_{s \in k} \overline{p}_j^s(\pi)) = -v_j \tag{39}$$

If we restrict the model to telephone networks with load sharing strategy and without call admission control on the path level (all calls are accepted if there is free capacity), it can be shown that the average shadow prices in (39) are the unique solutions to the equations

$$\overline{p}^s(\pi^f) = \frac{\lambda^s}{\overline{\lambda^s}}(E(\lambda_j^s, N^s - 1) - E(\lambda_j^s, N^s)) \sum_i \overline{\lambda}_i^k r_i^s(\pi^f) \tag{40}$$

where $E(.)$ denotes the Erlang's formula, i is the index of stream using the s-th link with N^s trunks, λ^s, $\overline{\lambda^s}$ denote the rates of the superposition of all calls offered to and accepted on the link s, respectively. Almost identical results as (39, 40) were obtained by Kelly [17] for a telephone network with the load sharing policy but based on a different model (no MDP). In fact the only difference is in the definition of the link reward parameter which in [17] is given by (19) in place of (17). The explanation of this difference is that although in both cases the evaluation of the average shadow prices (shadow price in [17]) is based on the decomposition of the Markov process, resulting in the Erlang fixed-point approximation, the reward process is treated differently. In the case of MDPD the reward process is also decomposed providing that $r_j = \sum_{s \in k} r_j^s(\pi)$ and consequently $\overline{R}(\pi) = \sum_s \overline{R}^s(\pi)$. This feature can significantly simplify network dimensioning and the routing optimization problem with blocking constraints (see [23]). In the case of Kelly's model the reward process is not decomposed and it is easy to show that in general $r_j \neq \sum_{s \in k} r_j^s(\pi)$ so the separable link average rewards cannot be defined.

VI. MDPD Performance Analysis

The performance analysis was performed for the network examples described in Table III. The levels of traffic and network structures in examples W7N, N7N and N11N are similar to the examples based on the ITU data for telephone networks [19]. They are non symmetrical but well connected and well dimensioned. The examples W7S and N7S are fully connected and symmetrical. Finally the examples W8N, W6N, created to show some particular features, are not well connected and not well dimensioned (e.g. specific hour in the multihour case). The structure of examples W8N, W6N are presented in Fig.3. In all examples the length of alternative paths is limited to two links. The policy iteration procedure was implemented with the direct routing, π^d, as the initial policy and the value iteration algorithm was used to evaluate link shadow prices. The applied link model assumes that the filtering probability is equal 1 for all link states (model SAR from Section IV.*C.1*), uses D2 model for call reward parameter division (Section IV.*A*) and aggregates all link call classes with the same bandwidth requirements. The 95% confidence intervals for simulation results are presented in the tables. For the sake of the presentation clarity they are omitted in the figures, nevertheless they can be approximately assessed from the corresponding examples in tables.

Since the exact solution is not computable for any reasonable example, in the following we use an indirect analysis to investigate important features of the model.

TABLE III
Description of Network Examples

	W7N	W7S	W8N	W6N	N7N	N7S	N11N
symmetrical	no	yes	no	no	no	yes	no
# nodes	7	7	8	6	7	7	11
d_j	1, 6	1, 12	1, 12	1, 6	1	1	1
μ_j^{-1}	1, 10	1, 10	1, 10	1, 1	1	1	1
WB traffic	50 %	49 %	35 %	33 %	0 %	0 %	0 %
traffic [Erl.]	2466	2058	394	203	1137	1722	10355
overload [%]	+10	+10	+20	+20	+10	+10	+10
link capacity	0-200	120	120	120	7-170	50	0-439
$r_j' = r_j\mu_j/d_j$	1	1	1	1	1	1	1

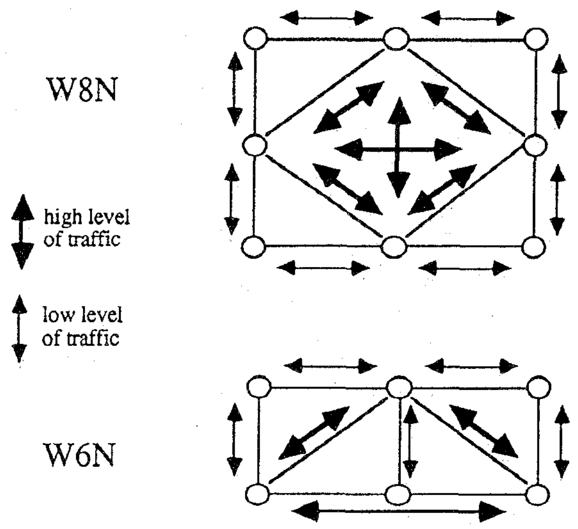

Fig. 3. Network connectivity and traffic structure for examples W8N and W6N.

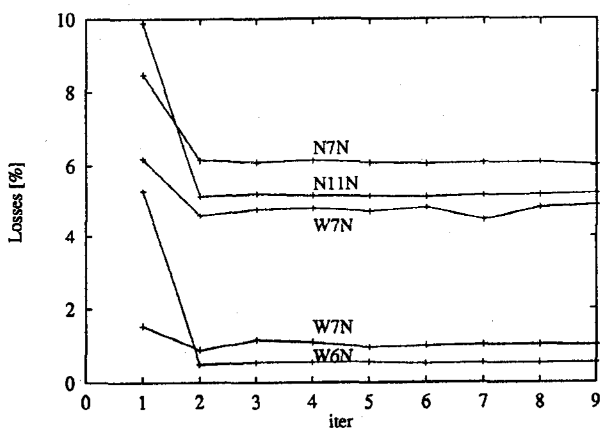

Fig. 4. Convergence of the policy iteration algorithm.

A. Convergence

The convergence of the proposed policy iteration scheme can be influenced by the accuracy of statistics used for evaluation of link arrival rates. To minimize influence of this factor in our study the policy updating period is equal to ten maximum mean holding times. Typical examples of the convergence of the algorithm are presented in Fig.4 where the average reward losses, H, are given as a function of the number of iteration cycles, i. Observe that the policy achieved after the first iteration (shadow price evaluation based on direct routing) is equivalent to the approach proposed in [14, 15] for telephone network (see Section III). The results indicate that by neglecting the multilink flows the performance of the routing can be deteriorated by several hundreds of percent (examples W6N and N11N). The convergence of the proposed algorithm is very fast and in all tested cases two iterations are sufficient to achieve policy close to the limit one. In the sequel all results are given for 8 iterations (arbitrary number).

B. Path Selection

To investigate the optimality of flow distribution we consider two modifications of the path selection algorithm. In the first one the direct link has priority. It means that a new call is offered to the direct link whenever the direct link net-gain is positive, otherwise a path with maximum positive net-gain is chosen. In fact the priority for direct link is commonly used in most existing and proposed routing schemes for circuit switched networks (e.g. [16,30]). The second modification utilizes the result from sensitivity analysis presented in Section V. Namely we increase probability of choosing path with higher value of $\frac{d\overline{R}}{d\overline{\lambda}_j^k}$, by modifying the path net-gain (for path selection purpose only) as follows

$$\begin{aligned} g_j^{k'}(\mathbf{y},\pi) &= (1-\alpha)g_j^k(\mathbf{y},\pi) + \alpha\frac{d\overline{R}}{d\overline{\lambda}_j^k} \\ &= r_j - \sum_{s\in k}[(1-\alpha)p_j^s(\mathbf{x}^s) + \alpha\overline{p}_j^s] \qquad (41) \end{aligned}$$

TABLE IV
IMPACT OF DIFFERENT PATH SELECTION SCHEMES ON REWARD LOSSES [%]

	MDPD $\alpha = 0$	priority $\alpha = 0$	MDPD' $\alpha = 0.6$	SEQ $\alpha = 1.0$
		nominal conditions		
W7N	1.03 ± .12	0.79 ± .14	0.83 ± .20	0.88 ± .15
W7S	1.10 ± .17	0.97 ± .23	0.90 ± .15	1.21 ± .27
W8N	0.91 ± .17	25.81 ± .37	0.70 ± .14	2.29 ± .22
N7N	1.42 ± .09	1.02 ± .09	1.07 ± .09	1.18 ± .10
N7S	0.76 ± .07	0.27 ± .03	0.26 ± .04	0.48 ± .05
N11N	0.51 ± .06	0.22 ± .03	0.20 ± .05	0.26 ± .05
		overload conditions		
W7N	4.83 ± .30	4.34 ± .24	4.57 ± .37	4.27 ± .30
W7S	5.54 ± .43	4.80 ± .40	5.13 ± .45	5.46 ± .44
W8N	5.02 ± .31	33.87 ± .25	4.92 ± .29	10.50 ± .26
N7N	6.10 ± .10	6.08 ± .11	6.00 ± .11	6.18 ± .10
N7S	4.60 ± .12	4.38 ± .12	4.42 ± .11	4.90 ± .10
N11N	5.19 ± .11	5.34 ± .12	5.11 ± .13	5.28 ± .10

where $\alpha = [0,1]$ is a weighting factor and $\frac{d\overline{R}}{d\overline{\lambda}_j^k}$ is substituted using (37). The rational of this modification can be explained as follows. Firstly, by increasing the flow on the paths with higher value of $\frac{d\overline{R}}{d\overline{\lambda}_j^k}$ the algorithm tries to equalize, whenever possible, the average net-gain of the used paths (the optimality condition for load sharing strategy). Secondly, the model for evaluation of the average shadow prices is not directly influenced by the link independence assumption thus the derivative $\frac{d\overline{R}}{d\overline{\lambda}_j^k}$ is less biased by this assumption (compared to the state dependent path net-gain).

The results presented in Table IV show that priority for direct link provides marginal improvement in the performance of well connected and well dimensioned network examples especially in nominal conditions. But in the case of W8N example the performance is significantly deteriorated by the modification. These results indicate that in the MDPD the flow distribution can be slightly biased by the link independence assumption. On the other hand it is clear that priority for direct link should not be used as a general solution.

The performance of the second modification as a function of α is presented in Fig.5. The performance for $\alpha = 0.6$ (MDPD') and $\alpha = 1.0$ (SEQ) is also presented in Table IV. In the first phase of increasing α the control performance is approaching the performance of the case with priority for direct link when this scheme provides performance improvement. Thus, in this range, the feature of an adaptive load sharing routing is correcting the flow distribution. In the second phase, when α approaches unity the average blocking is increasing in most cases. This is caused by the fact that in this range all paths for given OD are ordered according to the value of $\frac{d\overline{R}}{d\overline{\lambda}_j^k}$ and the sequential choice in many cases does not provide optimal flow dis-

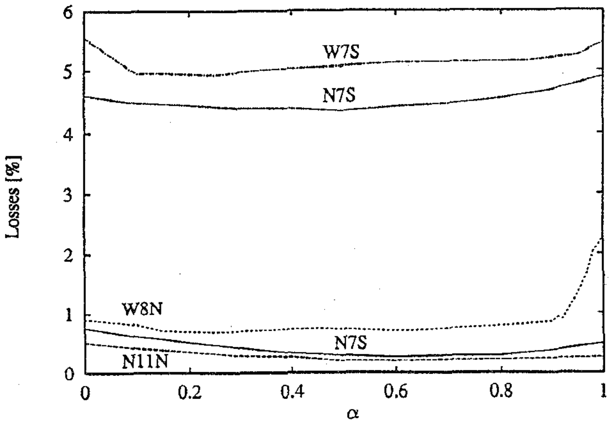

Fig. 5. **Reward losses vs. weighting factor α.**

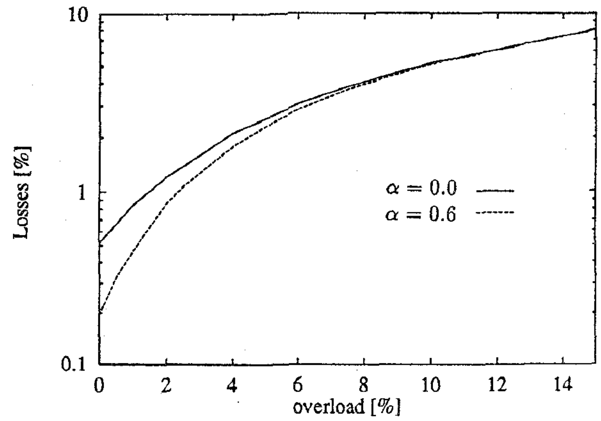

Fig. 7. **Reward losses vs. overload factor (ex. N11N).**

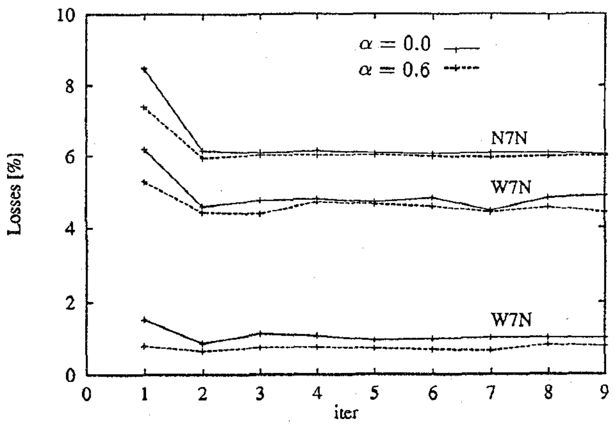

Fig. 6. **Convergence of the algorithm with $\alpha = 0.0$ and $\alpha = 0.6$.**

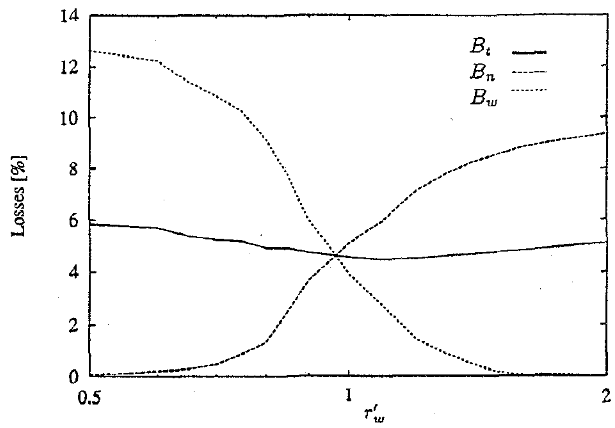

Fig. 8. **Traffic losses vs. reward parameter r'_w in network example W7N.**

tribution. From a practical point of view it is important that in most tested cases the optimal value of α falls in the interval $[0.3, 0.7]$ and that the function $H(\alpha)$ is shallow in this interval. This indicates that one value of α can provide close to optimal performance for all networks. Based on this premiss we use $\alpha = 0.6$ (MDPD') in all subsequent considerations.

The correction factor has also influence on the convergence of the algorithm. Namely besides improving performance of the final policy ($i \geq 2$) it can also significantly improve the performance of the policy after first iteration. This is illustrated in Fig.6. Finally it should be stated that under heavy overload the flow distribution error is disappearing since the probability that a multilink path is chosen when direct link is available is negligible. This is shown in Fig.7.

C. GOS Distribution Control

To illustrate the powerful influence of call reward parameters we present in Fig.8 the blocking probability of NB calls, B_n (average over all classes), the blocking probability of WB calls, B_w (average over all classes) and the weighted overall blocking probability, B_t, versus normalized reward parameter of WB calls, r'_w ($r'_j = r_j \mu_j / d_j$). As can be seen the reward parameters provide a tool for controlling the ratio of WB traffic losses to NB traffic losses over a very wide range including their equalization. The control of individual stream is illustrated in Fig.9. In this case the blocking of WB stream, B_j, offered to the j-th origin destination node pair can be controlled over a wide range by its reward parameter, r'_j, while the average network losses, B_t, are little influenced.

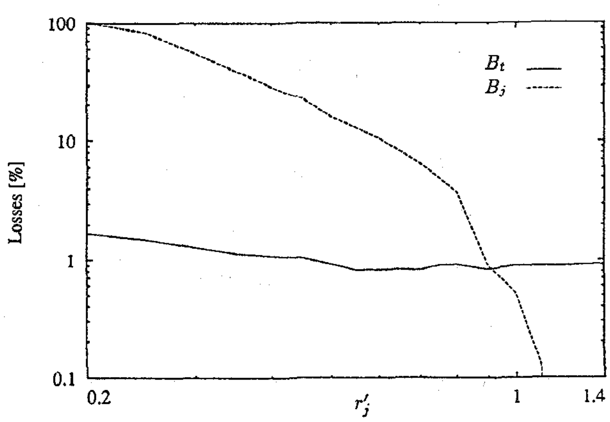

Fig. 9. **Traffic losses vs. reward parameter r'_j in network example W7S.**

TABLE V
REWARD LOSSES [%] FOR DIFFERENT ROUTING STRATEGIES

	MDPD'	LLP	DLM
		nominal conditions	
W7N	0.83 ± .20	1.20 ± .08	0.96 ± .16
W7S	0.90 ± .15	1.17 ± .08	0.89 ± .12
W8N	0.70 ± .14	9.93 ± .20	0.89 ± .30
N7N	1.07 ± .09	1.26 ± .09	
N7S	0.26 ± .04	0.37 ± .02	
N11N	0.20 ± .05	0.34 ± .07	
		overload conditions	
W7N	4.57 ± .37	5.71 ± .18	5.29 ± .28
W7S	5.13 ± .45	5.97 ± .29	5.31 ± .33
W8N	4.92 ± .29	13.15 ± .23	5.58 ± .41
N7N	6.00 ± .11	6.42 ± .07	
N7S	4.42 ± .11	4.53 ± .07	
N11N	5.11 ± .13	6.93 ± .03	

D. Comparison with LLP and DLM Strategies

For the comparison with the reward maximization approach we have developed a policy which is an extension of the least loaded path approach proposed in [30]. In the extended model a call from class i is offered to the direct link first and if it is blocked a recommended path is tried. The recommended path is chosen randomly with the probability proportional to the path's available capacity defined as:

$$\overline{x}_i^k = \min_{s \in k}[N^s - \sum_i x_i^s d_i - t_i^s - h_i^s(\mathbf{x})] \qquad (42)$$

where t_i^s denotes the trunk reservation level (protecting direct calls against multilink calls) and $h_i^s(\mathbf{x})$ denotes the number of of trunks reserved for WB calls in state $\mathbf{x}$ (protecting WB calls against NB calls). Concerning the evaluation of t_i^s we tried approaches which varied from the published adaptive schemes via some heuristics to a search through the space of all possible combinations in case of symmetrical networks. The best results are reported. Concerning the evaluation of $h_i^s(\mathbf{x})$ we applied the the dynamic bandwidth allocation algorithm, DBA [13, 27] (also used in the DLM strategy). It is clear that $h_i^s(\mathbf{x}) = 0$ for WB calls. As has been shown in [13, 27] the performance and structure of DBA are close to optimal. Note that in the DBA model the link bandwidth is shared among services. This is in contrast with [7] where a part of the bandwidth is dedicated to particular services and limits on the maximum number of each type calls are set.

The comparison of the performance of MDPD', LLP and the approach based on link model decomposition, DLM (Section IV.C), is given in Table V. In all tested examples the performance of the MDPD' model is better than that of the LLP model though the differences are relatively small in most cases. The exception is the example W8N where the performance of the LLP model is significantly worse. This is mainly caused by the priority for direct link scheme used in the LLP approach. Although the results are achieved for a particular version of the LLP approach the authors believe that the qualitative conclusions can be generalized to all LLP schemes due to some inherent features of the MDPD' and LLP schemes. Obviously the general argument could be that the MDPD' approach is derived from the optimal model. But this statement can be also supported by two more detailed yet simple arguments. Firstly, although both the MDPD' and the LLP schemes are using all link states to make the decision, in the case of LLP only the state of the path's link with smallest free capacity counts. It means that the probability of choosing the path is not changed for a whole range of states on the other links providing that their free capacity is larger than or equal to the bottleneck. In the case of MDPD', the state of each link influences the path net-gain thus more information about the network state is used in the decision. Secondly, the decision in the MDPD' model takes into account the current flow distribution in the network by using the traffic measurements in the evaluation of the shadow prices. This feature allows to model the system dynamics. In the case of LLP the network state information is static.

The performance of the DLM model is in all cases close to the MDPD' model. This result confirms the conclusion from Section IV.C that simplified link models for shadow price evaluation can provide good performance.

VII. CONCLUSIONS

The model presented in the paper contributes to the field of network control in several areas. From the application point of view the main contribution is synthesis of a control strategy for multiservice networks where all basic functions, including control of each service GOS, are included and optimized in one homogeneous, theoretically based model. The important feature of this approach is the real-time traffic measurements which feed the model with the current flow distribution and provides that the control policy adapts to a time variable traffic demand. It is also significant that once the scheme is implemented the control of GOS can be executed by simply varying the call reward parameters. From the modeling point of view the main contribution is the decomposition of the network reward process into separable link reward processes. This decomposition permits an implementable solution of the network Markov decision problem while the full policy iteration procedure is preserved. The approach, together with its sensitivity analysis provides also a general framework for studying, constructing and optimizing other call admission and routing strategies.

The proposed approach also opens up some new areas of investigation. In particular decomposition of the network reward process into a set of separable link network processes allows one to introduce new economical considerations in network dimensioning, planning and management. The basis for these considerations is setting a relation between the link cost, reward from the link and call reward parameters. A preliminary discussion of these issues is given in [23]. Another potential area open to investigation is application of the model to control ATM based networks. This might require further simplifications in the link model. Some preliminary results on this subject are given in [31]. Finally, a model for sensitivity of blocking probabilities with respect to the reward parameters would help to take full advantage of the ability to control GOS distribution.

Acknowledgment

The authors wish to thank Ke-Qiang Liao for his contribution concerning decomposition of link analysis and the simulation model. The authors are also grateful to Nicole Tetreault for her help in the numerical study. Finally we would like to acknowledge the effort of anonymous reviewers whose remarks have influenced the final shape of this paper.

References

[1] K.S. Narendra, E.A. Wright and L.G. Mason, "Application of Learning Automata to Telephone Traffic Routing and Control" IEEE Trans. Systems, Man and Cybernetics, Vol. SMC-7, No.11, Nov. 1977

[2] G.R. Ash, R.H. Cardwell and R.P. Murray, "Design and optimization of networks with Dynamic Routing", BSTJ, vol.60, no.8, October, 1981.

[3] C. Grandjean, "Call Routing Strategies in Telecommunication Networks", Proc. of ITC 5, New York, USA, 1967.

[4] E. Szybicki and M.E. Lavigne, "The introduction of an advanced routing system into local digital networks and its impact on networks' economy, reliability and grade of service", Proc. of ISS, Paris, France, 1979.

[5] G.R. Ash, "Use of a Trunk Status Map for Real-Time DNHR", Proc. of ITC 11, Kyoto, Japan, 1985.

[6] G.R. Ash, J.-S. Chen, A.E. Frey and B.D. Huang, "Real-Time Network Routing in a Dynamic Class-of-Service Network", Proc. of ITC 13, Copenhagen, Denmark, 1991.

[7] G.R. Ash, B.M. Blake and S.D. Schwartz, "Integrated network routing and design" Proc. of ITC 12, Torino, Italy, 1988.

[8] I.S. Gopal and T.E. Stern, "Optimal Call Blocking Policies in an Integrated Services Environment" Conf. Inform. Sci. Syst., The Johns Hopkins Univ., 1983, pp. 383-388.

[9] K. Miyake, "Optimal Trunk Reservation Control for Multi-slot Connection Circuits", Trans. of IEICE, November 1988.

[10] T. Oda and Y. Watanabe, "Optimal Trunk Reservation for a Group with Multislot Traffic Stream", IEEE Trans. Commun., vol.38, No.7, July 1990.

[11] K. Lindberger, "Blocking for Multislot Heterogeneous Traffic Streams Offered to a Trunk Group with Multislot Traffic Streams", Proc. of 5th ITC Specialist Seminar, Lake Como, Italy, May 1987.

[12] A. Gersht and K.J. Lee, "Virtual-Circuit Load Control in Fast Packet-Switched Broadband Networks" Proc. of IEEE GLOBECOM'88, Hollywood, Florida, USA, 1988.

[13] K.-Q. Liao, Z. Dziong and L.G. Mason, "Dynamic Link Bandwidth Allocation in an Integrated Services Network", Proc. of IEEE ICC'89, Boston, USA, June 1989.

[14] W.G. Lazarev and S.M. Starobinets, "The use of dynamic programming for optimization of control in networks of commutations of channels", Engineering Cybernetics (Academy of Sciences, USSR), No.3, 1977.

[15] K.R. Krishnan and T.J. Ott, "State dependent routing for telephone traffic: theory and results", Proc. of 25th IEEE Conference on Decision and Control, Athens, Greece, December 1986.

[16] K.R. Krishnan and T.J. Ott, "Forward-Looking Routing: A New State-Dependent Routing Scheme", Proc. of ITC 12, Torino, Italy, 1988.

[17] F.P. Kelly, "Routing in circuit switched networks: optimization, shadow prices and decentralization", Adv. Appl. Prob. 20, 1988.

[18] E. Szybicki, "Adaptive, Tariff Dependent Traffic Routing and Network Management in Multi-Service Networks", Proc. of ITC 11, Kyoto, Japan, 1985.

[19] Z. Dziong, M. Pioro, U. Körner and T. Wickberg, "On Adaptive Call Routing Strategies in Circuit Switched Networks - Maximum Revenue Approach ", Proc. of ITC 12, Torino, Italy, 1988.

[20] R. A. Howard, "Dynamic Programming and Markov Process", The M.I.T. Press, Cambridge, Massachusetts 1960.

[21] H. C. Tijms, "Stochastic Modeling and Analysis: A Computational Approach" New York: Wiley, 1986.

[22] K.W. Ross and D.H.K. Tsang, "Optimal Circuit Access Policies in an ISDN Environment: A Markov Decision Approach", IEEE Trans. Commun., vol.37, pp. 934- 939, September 1989.

[23] Z. Dziong and K.-Q. Liao, "Reward Maximization as a Common Basis for Routing, Management and Planning in ISDN", Proc. of NETWORKS'89, Spain, September 1989.

[24] A. Girard, "Routing and Dimensioning in Circuit-Switched Networks", Addison-Wesley, 1990.

[25] K.R. Krishnan, "Performance Evaluation of Networks Under State-Dependent Routing", Proc. of Seminar on Design and Control of a Worldwide Intelligent Network, Paris, June 1990.

[26] Z. Dziong and L.G. Mason, " Control of Multi-Service Loss Networks", Proc. of The 28th IEEE Conference on Decision and Control, Tampa, Florida, USA, December 1989.

[27] Z. Dziong and L.G. Mason, "An Analysis of Near Optimal Call Admission and Routing Model for Multi-Service Loss Networks", Proc. of INFOCOM'92, Florence, Italy, May 1992.

[28] M.I. Reiman and B. Simon, " Open Queuing Systems in Light Traffic" Math. of Oper. Res., 14, 1, 26-59, 1989.

[29] P. Bremaud, F. Vazquez-Abad, "On the pathwise estimation of derivatives w.r.t. the rate of a Poisson process, the phantom RPA method", QUESTA, 1992.

[30] W.H. Cameron, P. Galloy and W.J. Graham, "Report on the Toronto Advanced Routing Concept Trial", Proc. of Telecommunication Networks Planning Conference, Paris, France, 1980

[31] Z. Dziong, K-Q. Liao, L.G. Mason and N. Tetreault, "Bandwidth Management in ATM Networks", Proc. of ITC 13, Copenhagen, Denmark, June 1991.

Zbigniew Dziong received his M.Sc. and Ph.D. degrees from the Warsaw University of Technology, Poland, in 1974 and 1980, respectively, both in Electrical Engineering. From 1974 to 1987 he was with the Warsaw University of Technology as an Assistant Professor. He was on sabbatical leaves at the Centre National d'Études des Télécommunications, Paris, France, and at the Department of Communication Systems, Lund Institute of Technology, Sweden. Since 1987 he has been with INRS-Télécommunications, Canada, as a Professor. He also helds a position of Research Fellow at the Canadian Institute of Telecommunication Research. His research interests include design, management and control of communication systems in broadband networks. He was co-recipient of the 1993 STENTOR Award in recognition of his work on state-dependent routing.

Lorne G. Mason obtained the B.Sc. and Ph.D. degrees in mechanical engineering from the University of Saskatchewan in 1963 and 1972, respectively. He was with Bristol Aerojet in Winnipeg, Manitoba, from 1963 to 1965 involved in the design of the Black Brant rockets. He joined the British Columbia Telephone Co. as a traffic engineer in 1966 and again in 1972 as a consultant for planning digital networks. In 1973 he served as a consultant to Yale University, where he, in collaboration with Professor K.S. Narendra, pioneered the use of learning automata for adaptive routing in telecommunication networks. Between 1974 and 1977, he was with Bell-Northern Research where he developed planning tools and methods for digital network evolution and state-dependent routing. Since 1977 he has been with INRS-Télécommunications where he currently holds the title of full professor. He also holds a position of "Professeur Associé" at ENST (Télécom Paris). His primary research interests are in the application of control theory and operations research methods to telecommunication network design and management. He has held numerous industrial research contracts and NSERC strategic grants in the area of broadband network design and analysis and has numerous publications on the subject. Professor Mason is a participant in the Canadian Institute for telecommunications Research one of the Network of Centers of Excellence programs sponsored by the Canadian government, where he is project leader on broadband network control. He was co-recipient of the 1993 STENTOR Award for collaborative research in telecommunications for his contributions to state-dependent routing. He has also participated in various national and international technical committees dealing with broadband networks.