On the Performance of Dynamic Online QoS Routing Schemes*

Antonio Capone, Luigi Fratta, and Fabio Martignon

DEI, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy
{capone,fratta,martignon}@elet.polimi.it

Abstract. Several dynamic QoS routing techniques have been recently proposed for new IP networks based on label forwarding. However, no extensive performance evaluation and comparison is available in the literature. In this paper, after a short review of the major dynamic QoS routing schemes, we analyze and compare their performance referring to several networks scenarios. In order to set an absolute evaluation of the performance quality we have obtained the ideal performance of any routing scheme using a novel and flexible mathematical programming model that assumes the knowledge of arrival times and duration of the connections offered to the network. This model is based on an extension of the maximum multi-commodity flow problem. Being an integer linear programming model, its complexity is quite high and its evaluation is constrained to networks of limited size. To overcome the computational complexity we have defined an approximate model, based on the multi-class Erlang formula and the minimum multi-commodity cut problem, that provides an upper bound to the routing scheme performance. The performance presented in the paper has been obtained by simulation. From the comparison of the schemes considered it turns out that the Virtual Flow Deviation routing algorithm performs best and it almost reaches, in several scenarios, the ideal performance showing that no much gain is left for alternate new schemes.

1 Introduction

The current evolution of Internet architecture is towards service differentiation and Quality of Services (QoS) support [1]. In order to offer guaranteed end-to-end performance (as bounded delay, jitter or loss rate), it is necessary to introduce some sort of resource reservation mechanism and traffic control. With classical IP routing, however, when the resources are not available on the shortest path, the connection request is rejected even if sufficient resources exist on alternative paths.

With new label based forwarding mechanisms, such as MPLS (Multi Protocol Label Switching) [2] and GMPLS (Generalized MPLS) [3, 4], per flow path selection is possible and QoS parameters can be taken into account by routing algorithms. The goal of QoS routing schemes is to select a path for each traffic flow

* This work has been partially supported under the grant of MURST Tango project.

M. Ajmone Marsan et al. (Eds.): QoS-IP 2005, LNCS 3375, pp. 456–469, 2005.
© Springer-Verlag Berlin Heidelberg 2005

(micro-flows or aggregated-flows according to routing granularity) that satisfies quality constraints based on the actual available resources in the network.

The QoS requirement of a connection can be given as a set of constraints on link and paths. For instance, bandwidth constraints require that each link on the path has sufficient bandwidth to accommodate the connection.

The QoS routing algorithms proposed in the literature [5–11] can be classified into static or dynamic, and online (on demand) or offline (precomputed) [6]. Static algorithms use only network information that does not change in time, while dynamic algorithms use the current state of the network, such as available link capacity and blocking probability. In online routing algorithms, path requests are considered one by one, and usually previously routed connections cannot be re-routed. Offline routing does not allow new path route computation and it is usually adopted for permanent connections.

From the user point of view QoS routing algorithms must satisfy the QoS requirements, while from the provider point of view they have also to maximize the resource utilization. For online routing schemes the maximum resource utilization is achieved by minimizing connection rejection probability of future requests.

This paper is focused on the performance evaluation of dynamic online QoS routing algorithms [12].

First, we review some of the most popular algorithms proposed in the literature, such as the Min-Hop Algorithm (MHA) [13], the Widest Shortest Path Algorithm (WSP) [14], the Minimum Interference Routing Algorithm (MIRA) [15], the Profile-Based Routing algorithm (PBR) [11] and the Virtual Flow Deviation (VFD) algorithm [17]. We describe in some detail MIRA and VFD algorithms. These algorithms take explicitly into account the topological layout of the ingress and egress points of the network. The VFD algorithm, recently proposed in [17], considers also the traffic statistics. More precisely, VFD exploits the knowledge of the layout of the ingress/egress nodes of the network, and uses the statistics information about the traffic offered to the network in order to forecast future connections arrivals.

Then, to provide a measure of the quality of the performance, we present some theoretical bounds to the performance achievable by any online QoS routing algorithm by means of two novel and flexible mathematical models.

The first one, Ideal Routing (IR), is an Integer Linear Programming model and is based on an extension of the maximum multi-commodity flow problem [18]. It provides an optimal routing configuration capable of accommodating the traffic offered to the network. The model minimizes the number of rejected connections assuming that the connection arrival times and their durations are known. Accepted connections are provided a single path which is maintained for the whole connection lifetime (no re-routing is allowed). The IR model describes an ideal routing scheme that achieves the minimum connection rejection probability. However, due to the complexity of its formulation, the solution of this model requires long computing time and large memory, even with state of the art optimization tools [19]. Therefore, its applicability is limited to small size network scenarios.

The second model, based on the multi-class Erlang formula and on the minimum multi-commodity cut problem [20–22] (Min-Cut model), is an approximate one and provides a looser lower bound to the connection rejection probability. It can be applied to larger and more complex network topologies since its memory occupation and computing time are considerably lower than in the first model.

The numerical results on the performance of the algorithms considered have been obtained by simulating a set of relevant network scenarios. The comparison of these results with the bounds obtained with the IR and Min-Cut models shows that the VFD algorithm performs quite close to the ideal algorithm.

The paper is structured as follows: in Section 2 we address the QoS routing problem and we review some existing routing algorithms. In Section 3 we review the Virtual Flow Deviation algorithm, pointing out its innovating features. In Section 4 we illustrate the IR model, discussing the problem of setting the objective function parameters, and the Min-Cut model. In Section 5 we analyze and discuss the performance of online algorithms under a variety of network scenarios, comparing their performance to the theoretical bounds calculated using the mathematical models. Section 6 concludes the paper.

2 Dynamic Online QoS Routing Schemes

In this section we review some of the most relevant dynamic QoS algorithms proposed in the literature. In the following we assume that all the quality parameters requested by incoming connections can be controlled by defining an equivalent flow bandwidth as discussed in [23, 24]. This assumption allows us to focus only on bandwidth constraints.

Let a network be represented by a graph $G(N, A)$, where the nodes N represent routers and arcs A represent communication links, as shown in Figure 1.

The traffic enters the network at ingress nodes S_i and exits at egress nodes T_i. Each connection requires a path from S_i to T_i. The capacity C_{ij} and the actual flow F_{ij} are associated to each link (i, j). The residual bandwidth of link (i, j) is defined as $R_{ij} = C_{ij} - F_{ij}$.

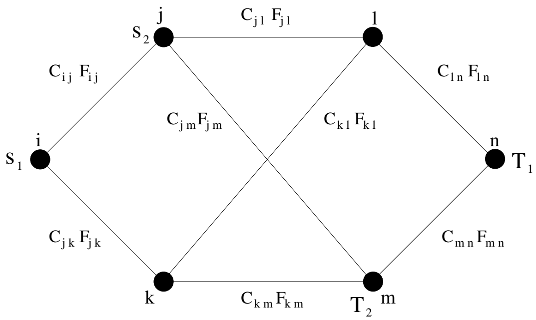


Fig. 1. QoS Network State.

A new connection can be routed only over links with R_{ij} greater or equal to the requested bandwidth. Referring to a new connection k with requested bandwidth d_k, a link is defined as *feasible* if $R_{ij} \geq d_k$. The feasible network for connection k is the sub-graph of G obtained by removing all un-feasible links. A connection can be accepted if at least one path between S_i and T_i exists in the feasible network. The minimum R_{ij} over a path defines the maximum residual bandwidth of that path.

The Virtual Flow Deviation (VFD) is a new routing algorithm, recently presented by the authors [17], that aims to overcome the limitations of the routing algorithms just reviewed by exploiting all the information available when a route selection must be taken.

To better describe the current state of the network and to forecast its future state, VFD exploits the topological information on the location of ingress/egress pairs, used by MIRA, as well as the traffic statistics obtained by measuring the load offered to the network at each source node. This information plays a key role in choosing the best route of a new request in order to prevent network congestion.

To account for the future traffic offered to the network, VFD routes not only the real call, but also some *virtual* calls which represent an estimate (based on measured traffic statistics) of the connection requests that are likely to interfere with the current real call. The number of these virtual calls, as well as the origin, destination, and the bandwidth requested should reflect as closely as possible the real future conditions of the network. These parameters can be estimated based on the past traffic statistics of the various ingress/egress pairs, as detailed in [17].

The accuracy of the measured traffic statistics is an important factor for the performance of the algorithm. However, even if the traffic statistics are not very accurate, the use of this information has shown to be effective and to provide better performance than simply using the topological information about the position of source and destination nodes as performed by MIRA.

All the information on the network topology and the estimated offered load is used to select a path which uses at best the network resources and minimizes the number of rejected calls. Such a path selection is performed in VFD by the Flow Deviation method [25, 26], which allows to determine the optimal routing for all connections entering the network.

3 Mathematical Models

In this Section we introduce two novel mathematical models that provide bounds to the performance achievable by any dynamic online routing algorithm.

The first model, Ideal Routing (IR), assumes the exact knowledge of future traffic. The routing decisions are taken to optimize the operation of the network loaded with the actual present and future traffic. No practical routing scheme can perform better. The model is based on an extension of the maximum multi-commodity problem and its solution obtained by ILP.

The second model, Min-Cut, releases all the constraints due to the network topology and optimizes the call acceptance assuming that the min-cut capacity of the network can be fully exploited. Its solution, based on minimum multi-commodity cut problem, is easier to obtain than IR. However, it provides a looser lower bound to the rejection probability that can anyway be used as a performance benchmark for large networks scenarios.

3.1 Ideal Routing Model

The basic assumption of the IR model is the knowledge of future traffic offered to the network. Let $K = \{1, \ldots, N_c\}$ be the set of connections, each one represented by the triplet (S_k, T_k, d_k) that specify source node, destination node and requested bandwidth. Connection k is further characterized by its arrival time, t_k and its duration τ_k. Given N_c connections (Fig. 2 shows an example for $N_c = 4$), the time interval from the arrival of the first connection and the last ending time of a connection is subdivided in a set I of $2N_c - 1$ time intervals.

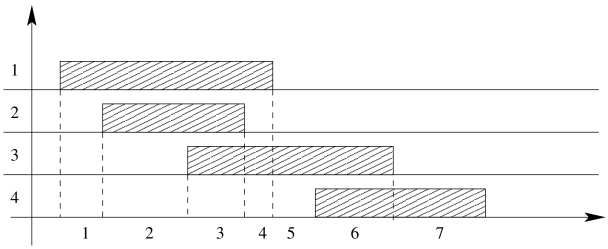


Fig. 2. Arrival time and duration of the connections offered to the network.

In each time interval, t, the number of active connections $M(t)$ remains constant. This number changes by one from interval to interval: it increases if a new connection arrives, and decreases if a connection ends. Let us denote with $B(k)$ the time interval beginning at the arrival time of connection k, and with I_k the set of time intervals in which connection k is active.

Given the function $M(t)$, the optimum routing must minimize the call rejection probability. This optimization problem can be formulated as Integer Linear Programming (ILP) if the following notations and definitions are adopted.

Let $G = (N, A)$ be the direct graph representing the network. Let $n = |N|$ and $m = |A|$ be the number of nodes and arcs, respectively. The capacity C_{ij} is associated to each arc (i, j).

For each connection k, $k \in K$, create two new nodes SS_k and TT_k and two new directed arcs, (SS_k, S_k) and (T_k, TT_k), of infinite capacity. Let SN and TN be the sets of the added nodes containing all SS_k and TT_k, respectively. Similarly, let A_{SN} and A_{TN} be the sets of arcs containing all (SS_k, S_k) and (T_k, TT_k), respectively.

Finally let $G' = (N', A')$ with $N' = N \cup SN \cup TN$ and $A' = A \cup A_{SN} \cup A_{TN}$.

Based on the above definitions and notation, we establish the ILP formulation of the IR model. To this purpose, let us define the following decision variables:

$$x_{ijt}^{k} = \begin{cases} 1 \text{ if connection } k \text{ is routed on arc } (i,j) \text{ in time slot } t \\ 0 \text{ otherwise} \end{cases}$$

for $(i,j) \in A'$, $k \in K$ and $t \in I$. We force $x_{ijt}^{k} = 0$, $\forall t \notin I_k$.

Since the goal is to minimize the connection rejection probability, we can equivalently maximize the number of connections accepted by the network. The problem can thus be formulated as follows:

$$Maximize \sum_{k \in K} b_k \cdot x_{SS_kS_kB(k)}^{k} \quad (1)$$

$$s.t. \sum_{k \in K} d_k \cdot x_{ijt}^{k} \leq C_{ij} \quad \forall (i,j) \in A, t \in I \quad (2)$$

$$\sum_{(j,l) \in A'} x_{jlt}^{k} - \sum_{(i,j) \in A'} x_{ijt}^{k} = \begin{cases} 1 \text{ if } j \in SN \\ 0 \text{ if } j \in N \\ -1 \text{ if } j \in TN \end{cases} \quad \forall k \in K, j \in N', t \in I \quad (3)$$

$$x_{ijt}^{k} = x_{ijB(k)}^{k} \quad \forall k \in K, (i,j) \in A', t \in I_k \quad (4)$$

$$x_{ijt}^{k} \in \{0,1\} \quad \forall k \in K, (i,j) \in A', t \in I_k \quad (5)$$

The objective function (1) is the weighted sum of the connections accepted in the network, where b_k represents the benefit associated with connection k. Different settings of b_k are possible, and they reflect different behaviors of the model as discussed later.

Constraints (2) ensure that, at each time slot, the total flow due to all the connections that use arc (i,j) does not exceed the arc capacity, C_{ij}, for all $(i,j) \in A$.

Constraints (3) represent the flow balance equations expressed for each node belonging to the extended graph G', in each time slot $t \in T$. Note that these constraints define a path for each connection between its source and destination nodes.

Constraints (4) impose that the accepted connections cannot be aborted or rerouted for their entire lifetime.

Finally, requiring that the decision variables in (5) are binary implies that each connection is routed on a single path.

The online QoS routing algorithms we are considering in this paper do not reject a new connection with (S_k, T_k, d_k) if at least one path with a residual available bandwidth greater than or equal to the requested bandwidth d_k exists.

To account for this feature, the objective function (1) must be properly set. To this purpose it is sufficient to set:

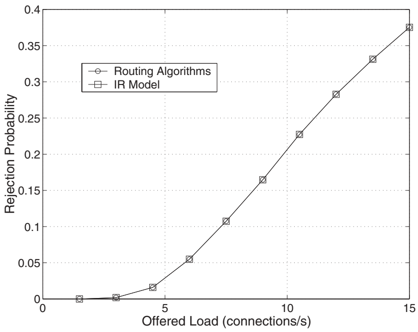
$$b_k = 2^{N_c - k} \quad (6)$$

having numbered the N_c connections from 1 to N_c according to their arrival times. With such a setting of b_k the benefit to accept connection k is always greater than the benefit of accepting, instead, all the connections from $k+1$ to N_c, since:

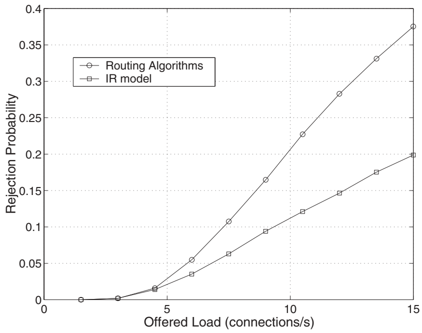
$$2^{N_c-k} > \sum_{i=k+1}^{N_c} 2^{N_c-i} \tag{7}$$

This choice of the weights b_k allows the mathematical formulation to model very closely the behavior of real online routing algorithms. To verify the accuracy of the model we have considered a simple scenario where a single link connects a source-destination pair. We have obtained the performance in the case of channel capacity equal to 20 bandwidth units and assuming the bandwidth b_k to be uniformly distributed between 1 and 3 units and the lifetime τ_k to be exponentially distributed with mean 15 s.

In this simple case all the routing algorithms provide the same performance since only one path exists between source and destination. The rejection probability shown in Fig. 3(a) has been computed using the multi-class Erlang Formula. The bound provided by the IR model completely overlaps the online routing performance. Note that different choices of b_k provide different IR Model performance. For instance, selecting $b_k = 1$ for all k we obtain the performance shown in Fig. 3(b). The large reduction in rejection probability is expected since the optimization of the objective function will result in rejecting connections with high bandwidth requirements and long lifetime in favor of smaller and shorter ones. The difference between the bound and the real performance increases as the network load increases.



(a)



(b)

Fig. 3. Connection rejection probability versus the average total load offered to a single link with (a) $b_k = 2^{N_c-k}$ (b) $b_k = 1$.

3.2 Min-Cut Model

In this Section we propose a second mathematical model that allows to determine a lower bound to the connection rejection probability. The solution of this model has computing times and memory occupation considerably lower than the previous one. However, in some scenarios, the bound obtained can be quite lower than the value provided by IR.

Let us consider a directed graph $G = (N, A)$ defined by a set of nodes, N, and a set of arcs, A each one characterized by a capacity C_{ij}. A set of source/destination pairs $K = \{1, ..., N_s\}$, indicated by S_i and T_i, respectively, $i \in K$, is also assigned. Each source S_i generates a flow αf_i, towards destination T_i, that can be split over multiple paths. The problem is to find the maximum α, indicated by α^*, such that for all $i \in K$ the flow quantities $\alpha^* f_i$ can be routed to their destinations.

The solution to this problem, obtained via linear programming techniques, provides the maximum multi-commodity flow $F_{max} = \sum_{i \in K} \alpha^* f_i$. Note that F_{max} represents a lower bound to the capacity of the minimum multi-commodity cut of the network, as discussed in [21, 22].

Once F_{max} has been obtained, the connection rejection probability for the given network scenario is obtained by using the multi-class Erlang formula with F_{max} servers [20] that is briefly reviewed in the following.

Let us consider N different traffic classes offered to a network system with C servers. The connections belonging to the class i request d_i bandwidth units. The connections arrival process is a Poisson process with average λ_i, while the connections duration is distributed according to a generic distribution $f_{\Theta_i}(\theta_i)$. Let $\Lambda = \sum_{i=1}^{N} \lambda_i$ be the total load offered to the network.

An appropriate state description of this system is $n = (n_1, ..., n_N)$, where $n_i, i = 1, ..., N$ is the number of connections belonging to the class i that occupy the servers. The set of all the possible states Ω is expressed as $\Omega = \{n | X \leq C\}$, where X, the total occupation of all the servers, is given by $X = \sum_{i=1}^{N} n_i d_i$.

If we indicate with $A_i = \lambda_i E[\Theta_i]$ the traffic offered to the network by each class, the steady state probability of each state is simply given by the multi-class Erlang formula:

$$\pi(n) = \frac{1}{G} \prod_{i=1}^{N} \frac{A_i^{n_i}}{n_i!} \tag{8}$$

where G is the normalization constant that ensures that the $\pi(n)$ sum to 1 and it has therefore the following expression:

$$G = \sum_{n \in \Omega} \pi(n) = \sum_{n \in \Omega} (\prod_{i=1}^{N} \frac{A_i^{n_i}}{n_i!}) \tag{9}$$

Using the steady state probability calculated with equation (8) we can derive the loss probability of the generic class i, Π_i, as follows:

$$\Pi_i = \sum_{n \in B_i} \pi(n) \tag{10}$$

where B_i is the set of the blocking states for the class i, defined as $B_i = \{n | C - d_i < X \leq C\}$. The overall connection rejection probability, p_{rej}, is then given by:

$$p_{rej} = \sum_{i=1}^{N} \frac{A_i \Pi_i}{\sum_{i=1}^{N} A_i} \tag{11}$$

If we substitute C with the maximum multi-commodity flow value F_{max} in all the above expressions, we can compute the connection rejection probability using equation (11).

In network topologies with high link capacities, F_{max} can assume high values, and the enumeration of all the allowed states becomes computationally infeasible, since the cardinality of Ω is of the order of F_{max}^N [27]. In these network scenarios, equations (8)-(11) are computationally too complex so we propose to apply the algorithm described in [27, 28] that computes recursively the blocking probability based on the peculiar properties of the normalization constant G. For network topologies with very high link capacities we implemented the inversion algorithm proposed in [29] to compute the blocking probabilities for each class.

4 Numerical Results

In this Section we compare the performance, measured by the percentage of rejected calls versus the average total load offered to the network, of the Virtual Flow Deviation algorithm, the Min-Hop Algorithm and MIRA with the bounds provided by the mathematical models presented in the previous Section referring to different network scenarios in order to cover a wide range of possible environments.

The first scenario we consider is illustrated in Figure 4. In this network the links are unidirectional with capacity equal to 120 bandwidth units. In the following capacities and flows are all given in bandwidth units. The network traffic, offered through the source nodes S_1, S_2 and S_3, is unbalanced since sources S_2 and S_3 generate a traffic four times larger than S_1. Each connection requires a bandwidth uniformly distributed between 1 and 3. The lifetime of the connections is assumed to be exponentially distributed with average equal to 15s.

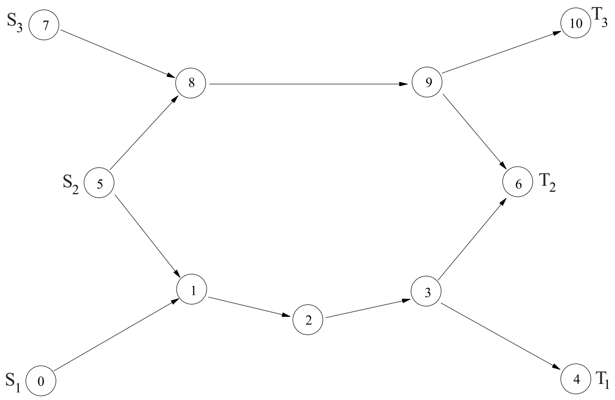


Fig. 4. Network topology with unbalanced offered load: the source/destination pairs S_2-T_2 and S_3-T_3 offer to the network a traffic load which is four times higher than that offered by the pair S_1-T_1.

In this simple topology connections S_1-T_1 and S_3-T_3 have one path only, while connections S_2-T_2 have two different paths.

The rejection probability versus the offered load for MIRA, MHA, VFD, IR and Min-Cut models are shown in Fig. 5. The poor performance of MIRA is due to its lack of considering any information about the load distribution in the network. In this particular topology, due to critical links (1,2), (2,3) and (8,9), S_2-T_2 connections are routed on the path (5-8-9-6) that contains the minimum number of critical links. MHA, that selects for connections S_2-T_2 the path with the minimum number of hops, routes the traffic as MIRA and their performances overlap. Better performance is achieved by VFD. Since its behavior depends on the number of virtual connection N'_v used in the routing phase, we have considered three cases: $N'_v = 0$, $N'_v = 0.5 \cdot N_v$ and $N_v = \lfloor (N_{max} - N_A) \rfloor$. In the first case, even if no information on network traffic statistics is taken into account, the VFD algorithm achieves much better performance than previous schemes due to the better traffic balance provided by the Flow Deviation algorithm. Only when the offered load reaches very high values the improvement reduces. The third case corresponds to the VFD version described in Section 3.2 that takes most advantage from traffic information. The best performance has been measured and the gain over existing algorithms is provided even at high loads. An intermediate value of N'_v (case 2) provides, as expected, intermediate performance. As far as the performance of the two mathematical models, we observe that the approximate Min-Cut model curve overlaps that of the IR model. Note that VFD performs very close to the theoretical bounds in this scenario.

To investigate the impact of connection lifetime distribution, we have considered a Pareto distribution with the same average as the previous exponential distribution and several shape parameters ($\alpha = 1.9, 1.95, 2.1, 3$). The performance observed in all cases are within 1% of those shown in Fig. 5.

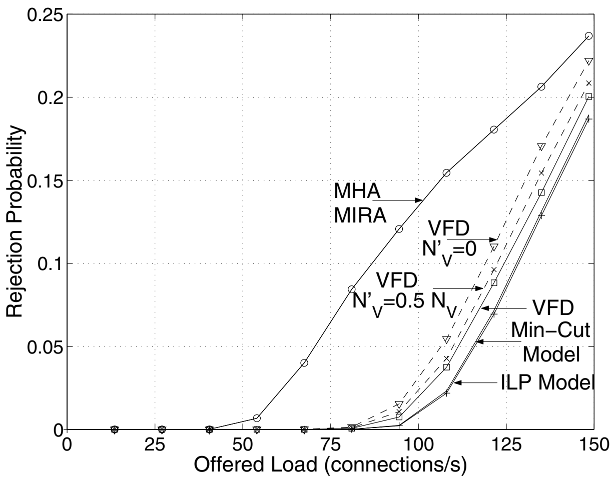
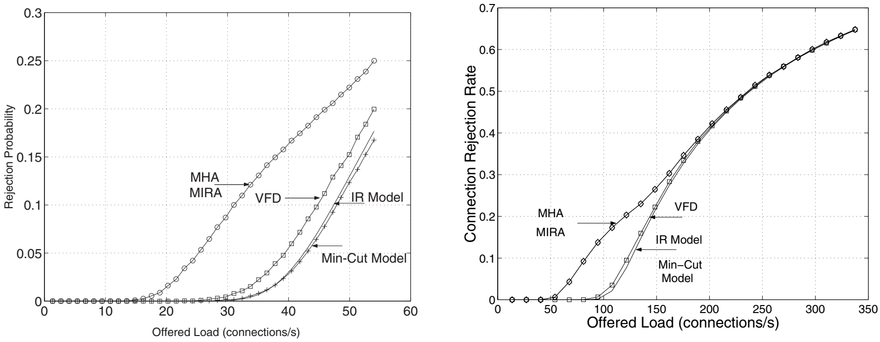


Fig. 5. Connection rejection probability versus the average total load offered to the network of Figure 4.

To test the sensitivity of the performance to the network capacity, we have considered, for the network in Fig. 4, different parameters. The results, shown in Fig. 6, are very similar to those of Fig. 5. It is worthwhile to observe that in all the different scenarios considered the approximate model provides results very close to IR. This validates the approximate model that can be easily evaluated even in more complex networks.



(a) (b)

Fig. 6. Connection rejection probability versus the average total load offered to the network of Fig. 4 (a) with link capacity equal to 24 and bandwidth requests always equal to 1 (b) with link capacity equal to 60 and bandwidth requests always equal to 1.

A more realistic scenario that was first proposed in [15] is shown in Fig. 7. The links marked by heavy solid lines have a capacity of 480 while the others have a capacity equal to 120, in order to replicate the ratio between OC-48 and OC-12 links. The performance for the case of balanced offered traffic, considered in [15], are shown in Figure 8(a).

VFD and MIRA achieve almost the same performance and are much better than MHA. VFD presents a slight advantage at low load since it starts rejecting connections at an offered load 10% higher than MIRA. We have measured that a rejection probability of 10^{-4} is reached at an offered load of 420 connections/s by MIRA as opposed to 450 connections/s for VFD. Also in this case the IR model is computationally too demanding. Therefore, we applied the Min-Cut model with the inversion algorithm proposed in [29], as the maximum multicommodity flow is equal to 1200 bandwidth units.

If we consider on the same topology an unbalanced load where for instance traffic S_1-T_1 is four times the traffic of the other sources, the improvement in the performance obtained by VFD is much more significant. The results shown in Figure 8(b) confirm that unbalanced situations are more demanding on network resources and the rejection probability for the same given offered load is much higher. In these more critical network operations VFD has proved to be more effective providing improvements of the order of 20% and well approaching the lower bound provided by the Min-Cut model.

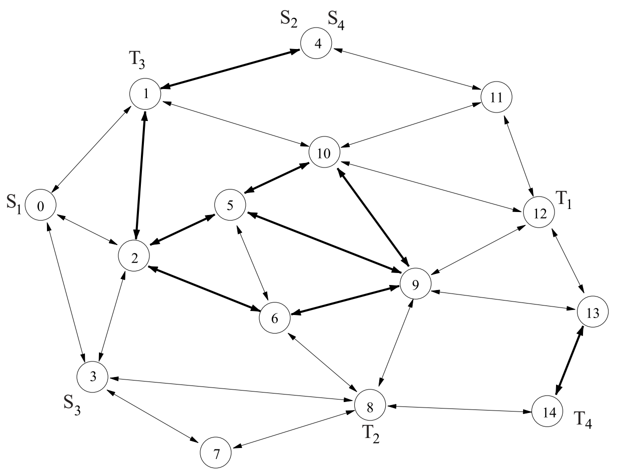
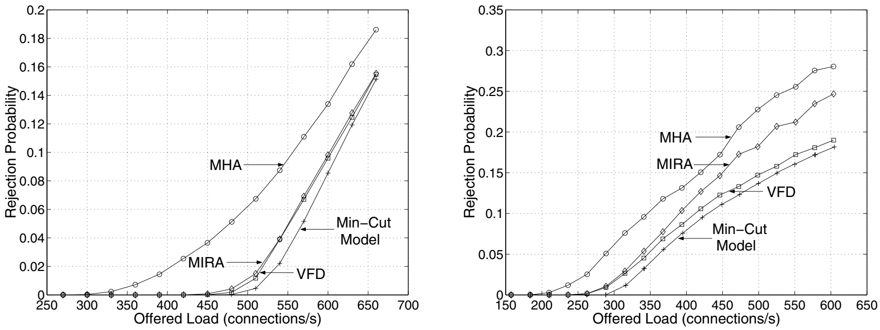


Fig. 7. Network Topology with a large number of nodes, links, and source/destination pairs.



(a) (b)

Fig. 8. Connection rejection probability versus the average total load offered to the network of Fig. 7 (a) where all sources produce the same amount of traffic; (b) where the traffic between S_1-T_1 is four times higher than the traffic produced by the other pairs.

5 Conclusions

We have discussed and analyzed the performance of online QoS routing algorithms for bandwidth guaranteed connections in MPLS and label switched networks.

To provide a theoretical bound on the performance achievable by dynamic online QoS routing algorithms we have proposed two novel mathematical models. The first is an Integer Linear Programming model that extends the well known maximum multi-commodity flow problem to include connections arrival-times and durations, while the second, which has a much lower complexity, is based on the application of the multi-class Erlang formula to a link with capacity equal to the residual capacity of the minimum network multi-commodity cut.

We have shown that the Virtual Flow Deviation scheme not only allows to reduce remarkably the blocking probability with respect to previously proposed routing schemes, but it also well approaches the lower bounds provided by the mathematical models in the considered network scenarios.

References

1. Hui-Lan Lu and I.Faynberg. An Architectural Framework for Support of Quality of Service in Packet Networks. *IEEE Communications Magazine*, pages 98-105, Volume 41, Issue 6, June 2003.
2. E.Rosen, A.Viswanathan, and R.Callon. Multiprotocol Label Switching Architecture. In *IETF RFC 3031*, January 2001.
3. Ed. L. Berger. Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description. In *IETF RFC 3471*, January 2003.
4. A.Banerjee, L.Drake, L.Lang, B.Turner, D.Awduche, L.Berger, K.Kompella, and Y.Rekhter. Generalized Multiprotocol Label Switching: an Overview of Signaling Enhancements and Recovery Techniques. *IEEE Communications Magazine*, pages 144 - 151, Volume 39 , Issue 7, July 2001.
5. S.Chen and K.Nahrstedt. An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions. *IEEE Network*, pages 64-79, Volume 12, Issue 6, November/December 1998.
6. J.L.Marzo, E.Calle, C.Scoglio, and T.Anjali. QoS Online Routing and MPLS MultilevelProtection: A Survey. *IEEE Communications Magazine*, pages 126 - 132, Volume 41 , Issue 10, October 2003.
7. Bin Wang, Xu Su, and C.L.P.Chen. A New Bandwidth Guaranteed Routing Algorithm for MPLS Traffic Engineering. In *IEEE International Conference on Communications, ICC 2002, Volume 2*, pages 1001–1005, April 2002.
8. Z.Wang and J.Crowcroft. QoS Routing for Supporting Resource Reservation. In *IEEE Journal on Selected Areas in Communications*, Sept.1996.
9. A.Orda. Routing with End to End QoS Guarantees in Broadband Networks. In *IEEE INFOCOM'98*, Mar.1998.
10. B.Awerbuch et al. Throughput Competitive On Line Routing. In *34th Annual Symp. Foundations of Computer Science*, Palo Alto, CA, Nov.1993.
11. S.Suri, M.Waldvogel, D.Bauer, and P.R.Warkhede. Profile-Based Routing and Traffing Engineering. *Computer Communications*, pages 351 - 365, Volume 26 , Issue 4, March 2003.
12. A.Capone, L.Fratta, and F.Martignon. Dynamic Online QoS Routing Schemes: Performance and Bounds. *submitted to Computer Communications*.
13. D.O.Awduche, L.Berger, D.Gan, T.Li, V.Srinivasan, and G.Swallow. RSVP-TE: Extensions to RSVP for LSP Tunnels. In *IETF RFC 3209*, December 2001.
14. R.Guerin, D.Williams, and A.Orda. QoS Routing Mechanisms and OSPF Extensions. In *Proceedings of Globecom*, 1997.
15. Murali S. Kodialam and T. V. Lakshman. Minimum Interference Routing with Applications to MPLS Traffic Engineering. In *Proceedings of INFOCOM*, pages 884–893, 2000.

16. Koushik Kar, Murali Kodialam, and T.V. Lakshman. Minimum interference routing of bandwidth guaranteed tunnels with MPLS traffic engineering applications. *IEEE Journal on Selected Areas in Communications*, December 2000, Volume 18, Issue 12.
17. A.Capone, L.Fratta, and F.Martignon. Dynamic Routing of Bandwidth Guaranteed Connections in MPLS Networks. *International Journal on Wireless and Optical Communications*, pages: 75-86, Volume 1, Issue 1, July 2003.
18. R.K.Ahuja, T.L.Magnanti, and J.B.Orlin. *Network Flows*. Prentice-Hall, 1993.
19. *ILOG CPLEX*. Available at http://www.ilog.com/products/cplex/.
20. J.F.P. Labourdette and G.W. Hart. Blocking probabilities in multitraffic loss systems: insensitivity, asymptotic behavior, and approximations. *IEEE Trans. on Communications*, 40:1355–1366, 1992.
21. Tom Leighton and Satish Rao. Multicommodity Max-Flow Min-Cut Theorems and their Use in Designing Approximation Algorithms. *Journal of the ACM*, Volume 46, Issue 6, November 1999.
22. Y.Aumann and Y.Rabani. An O(log k) Approximate Min-Cut Max-Flow Theorem and Approximation Algorithm. In *SIAM Journal on Computing*, pages 291–301, Volume 27, Issue 1, February 1998.
23. R.Guerin, H.Ahmadi, and M.Naghshineh. Equivalent capacity and its application to bandwidth allocation in high speed networks. In *IEEE Journal on Selected Areas in Communications*, pages 968–981, September 1991.
24. J.A.Schormans, J.Pitts, K.Williams, and L.Cuthbert. Equivalent capacity for on/off sources in ATM. In *Electronic Letters*, pages 1740–1741, Volume 30, Issue 21, 13 October 1994.
25. L.Fratta, M.Gerla, and L.Kleinrock. The Flow Deviation Method: An Approach to Store-and-forward Network Design. In *Networks 3*, pages 97–133, 1973.
26. D.Bertsekas and R.Gallager. *Data Networks*. Prentice-Hall, 1987.
27. J.S.Kaufman. Blocking in a Shared Resource Environment. *IEEE Transactions on Communications*, pages 1474 - 1481, Volume 29, Issue 10, 1981.
28. A.A.Nilsson, M.Perry, A.Gersht, and V.B.Iversen. On Multi-rate Erlang-B Computations. *in 16th International Teletraffic Congress; ITC'16*, pages 1051 - 1060, Elsevier Science, 1999.
29. G.L.Choudhury, K.K.Leung, and W.Whitt. An Inversion Algorithm to Compute Blocking Probabilities in Loss Networks with State-Dependent Rates. *IEEE/ACM Transactions on Networking*, pages 585 - 601, Volume 3, Issue 5, 1995.