Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Networks 50 (2006) 966–981
Computer Networks
www.elsevier.com/locate/comnet

Dynamic online QoS routing schemes: Performance and bounds

Antonio Capone [a,*], Luigi Fratta [a], Fabio Martignon [b]

[a] Department of Electronics and Information (DEI), Politecnico di Milano, Piazza Leonardo da Vinci 32, Room 330, Third Floor, 20133 Milan, Italy
[b] Department of Management and Information Technology, University of Bergamo, 24044 Dalmine, Italy

Received 14 August 2004; received in revised form 6 May 2005; accepted 20 June 2005
Available online 8 August 2005

Responsible Editor: H.J. Stuettgen

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, evaluated by measuring the connection rejection probability, shows that the schemes considered reach, in several scenarios, the ideal performance, showing that no much gain is left for alternate new schemes.

Keywords: QoS Routing; Mathematical models; MPLS; Flow deviation

* Corresponding author. Tel.: +39 02 2399 3449; fax: +39 02 2399 3413.
E-mail addresses: capone@elet.polimi.it (A. Capone), fratta@elet.polimi.it (L. Fratta), fabio.martignon@unibg.it (F. Martignon).

1389-1286/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2005.06.007

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 Multi Protocol Label Switching (MPLS) [2] and Generalized MPLS (GMPLS) [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 (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 links and paths. For instance, bandwidth constraints require that each link on the path has sufficient bandwidth to accommodate the connection.

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.

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. In online routing algorithms, path requests are considered one by one, and usually previously routed connections cannot be rerouted. Offline routing does not allow new path route computation and it is usually adopted for permanent connections.

This paper is focused on the performance evaluation of dynamic online QoS routing algorithms, where the maximum resource utilization is achieved by minimizing connection rejection probability of future requests.

First, we review some of the most popular algorithms proposed in the literature, such as the Min-Hop Algorithm (MHA) [12], the Widest Shortest Path Algorithm (WSP) [13], the Minimum Interference Routing Algorithm (MIRA) [14,15], the Profile-Based Routing algorithm (PBR) [11] and the Virtual Flow Deviation (VFD) algorithm [16]. We describe in some detail MIRA, PBR 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 [16], 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 [17]. 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 [18,19]. 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