

Profit-oriented resource allocation using online scheduling in flexible heterogeneous networks**

Peng Xu · George Michailidis · Michael Devetsikiotis

© Springer Science + Business Media, LLC 2006

Abstract In this paper, we discuss a generalized measurement-based adaptive scheduling framework for dynamic resource allocation in flexible heterogeneous networks, in order to ensure efficient service level performance under inherently variable traffic conditions. We formulate our generalized optimization model based on the notion of a “profit center” with an arbitrary number of service classes, nonlinear revenue and cost functions and general performance constraints. Subsequently, and under the assumption of a linear pricing model and average queue delay requirements, we develop a fast, low complexity algorithm for online dynamic resource allocation, and examine its properties. Finally, the proposed scheme is validated through an extensive simulation study.

Keywords Resource allocation · On line scheduling · Profit center

1. Introduction

Emerging networks are very large-scale, distributed systems, highly diverse in structure and topology, and exhibiting complex workload characteristics. Examples include distributed computing (and more general “grids”), fused *ad hoc* wired and wireless networks, broadband residential

**Portions of this paper have been presented at the 39th Conference on Information Sciences and Systems, CISS 2005, and at the Third IFIP Workshop on Next Generation Networks: Architectures, Protocols, Performance, 2005.

P. Xu · M. Devetsikiotis
Dept. of Electrical and Computer Engineering,
North Carolina State University,
Raleigh, NC 27695-7911, USA
e-mail: {pxu3, mdevets}@ncsu.edu

G. Michailidis
Department of Statistics, The University of Michigan,
Ann Arbor, MI 48109, USA
e-mail: gmichail@umich.edu

access systems, and 4th generation cellular networks that work in conjunction with their wired counterparts.

The value derived from and the efficient operation of such networks will hinge upon their instant availability and their agility to deliver a wide range of information-based and computing services, under varying conditions, and at the required level of quality. Such requirements ask for a *flexible* networking infrastructure regarding its monitoring and control aspects. Such non-homogeneous networks operating in a randomly changing environment should be characterized by their *configurable* capabilities in terms of *network environment awareness* and *adaptive controllability*.

Furthermore, bandwidth and delay-sensitive applications such as voice over IP, video-conferencing, online gaming, and interactive television, have made imperative the development of scheduling algorithms that provide differentiated service guarantees to multiple classes of traffic, and that accomplish this in a practical, implementable and efficient manner.

These emerging trends imply that *static* bandwidth reservation protocols accompanied by over-provisioning of access bandwidth and network links could lead to significant under-utilization of available resources. For such settings, a *dynamic* allocation of bandwidth (or network resources in general) that closely tracks the prevailing traffic characteristics and resource requirements can achieve significant savings, while at the same time satisfying service level guarantees. Implementation of such dynamic schemes requires *efficient traffic monitoring* and estimation policies coupled with adaptive bandwidth allocation mechanisms.

In this paper, we investigate the problem of online scheduling for flexible networks and access nodes. Specifically, we model a network node (e.g., access gateway) as a “profit center”, in which the price for different classes of traffic is predefined and the service guarantee constraints are also included in the general optimization model. This allows us to pose the resource allocation problem as one of profit maximization, subject to satisfying the service guarantee constraints imposed by the user classes.

Optimal resource allocation problems subject to service guarantee constraints have been investigated in various contexts in [9,8,12,16,4]. Specifically, Shin et al. proposed adaptive weighted-packet scheduling for premium service in Diff-Serv [16], in which weight are updated by the estimation of average queue size from exponential weighted moving average. On the other hand, Chandra et al. proposed the minimization of the “discontent” caused by delays, and used numerical methods to obtain the optimal solution [4]. However, the *simultaneous* consideration of a resource allocation scheme that utilizes online measurements and is based on a generalized pricing mechanism has not been studied in the literature before. Also, given the dynamic nature of service guarantee constraints, our profit-oriented optimization and its analytical investigation provide deeper insights and help shed light into the fundamental workings of such mechanisms.

The main methodological contributions of this paper are: First, we introduce a measurement-based adaptive scheduling framework for adaptive bandwidth allocation in flexible network nodes; Second, we formulate a generalized *profit center* optimization model for a flexible node that takes into consideration a pricing model for the various classes of users, as well as service guarantee constraints, such as packet loss rates and queueing delays; Third, given a generalized service model [13], we consider a linear pricing model subject to average queueing delay constraints and study the properties of the optimal solution obtained from a fast, low complexity algorithm. Fourth, the analysis of the allocation scheme is considered in both the over- (system capacity exceeds demand) and under-provisioned cases. Finally, the proposed scheme is validated through an extensive simulation study that considers several traffic scenarios.

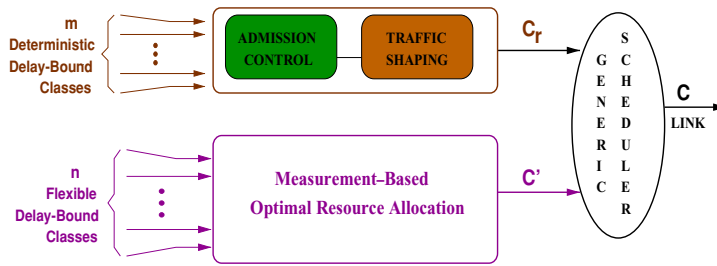


Fig. 1 General proposed framework to deliver service guarantees for flexible multi-class networks.

2. Modeling framework

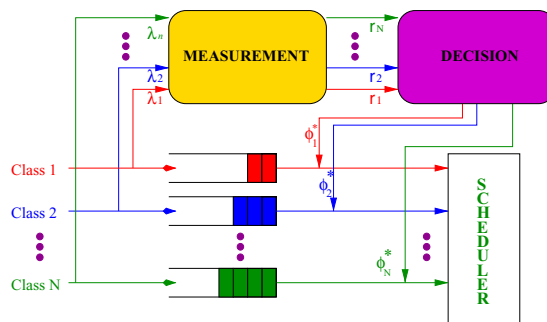
Our proposed modeling framework is illustrated in Fig. 1 for a single network element (switch or router). Suppose m deterministic delay-bound classes and n flexible delay-bound classes share the link capacity, C . Note that due to the extreme rigidity of the deterministic delay-bound classes, such service level requirements can only be guaranteed by the appropriate traffic shaping and admission control schemes ([6,9,11]), together with a certain amount of reserved bandwidth, C_r . Hence, these classes are excluded from our proposed Measurement Based Optimal Resource Allocation (MBORA) system.

A MBORA system is responsible of optimally allocating the excess bandwidth, $C' = C - C_r$, shared by the n flexible delay-bound classes. The proper allocation between the reserved bandwidth and the excess bandwidth can be achieved by a generic scheduler, such as Hierarchical Packet Fair Queueing mechanism proposed by Bennett and Zhang [2].

Next, we focus on the main components and related coordinations of a MBORA system shown in Fig. 2. Its main components are: a *traffic measurement module* that provides an accurate estimation of the future traffic load of the different classes under consideration over a pre-specified time interval (window), a *decision module* that determines how bandwidth is distributed among the various classes of traffic based on the information acquired from *traffic measurement module* and a *scheduling module* that deals with the packet forwarding mechanism.

The coordination of these three components is as follows: when a job/customer of class i arrives at the scheduler, it is assigned to the corresponding queue, waiting to receive service

Fig. 2 Illustration of the adaptive framework under consideration and its components.



from the scheduling module. At the same time, the measurement module updates the arrival rate statistics of the corresponding traffic class, and provides an estimate of the arrival rate. It should be noted that the measurement module performs the above operation over a pre-specified time interval (W window). Finally, the decision module allocates the service rate (bandwidth) to the queues using information about the estimated arrival rate and the queue length processes. In this paper, we focus on the problem that the decision module solves at every decision time instant (at the end of a W -duration window). Regarding the scheduling module, it updates the weights for the different classes of the generalized scheduler (WRR or WFQ), in order to accommodate the service rates determined by the decision module.

3. Generalized profit center model for resource allocation

3.1. Formulation

In this section, we formulate the bandwidth allocation problem associated with the decision module, that corresponds to a profit optimization problem solved by the network element. Hence, we take a social point of view [15], in allocating the available resources.

The provider's long-term profit consists of a revenue component based on usage charge, $\sum R_i(\phi_i)$, and a cost component related to the delay congestion, $\sum \mathcal{C}'_i(\phi_i)$, for all the classes, where ϕ_i is the proportion of the excess bandwidth allocated to class i . The objective is to maximize the usage charge while minimizing the delay penalty cost incurred by the congestion. Thus, our generalized profit center model can be formulated as the following optimization problem:

$$\max_{\phi} \sum_i U_i(\phi_i) = \sum_i R_i(\phi_i) - \sum_i \mathcal{C}'_i(\phi_i) \quad (1)$$

s.t.

$$\sum_{i=1}^N \phi_i \leq 1, \quad 0 \leq \phi_i \leq 1$$

$$P_{\text{loss}}^i(\phi_i) \leq P_{\text{loss}}^i$$

$$D_i(\phi_i) \leq d_i$$

where $P_{\text{loss}}^i(\phi_i)$ and $D_i(\phi_i)$ are class i loss rate and expected delay and P_{loss}^i, d_i are the pre-determined desired loss rate and loss probability, respectively. The first two sets of constraints correspond to the feasibility ones. It can be seen that the formulation of the profit center model depends on the characteristics of the pricing models adopted and on the nature of service guarantees and other constraints.

3.2. Discussion of pricing models and service guarantee constraints

The pricing model can have a linear or a non-linear functional form, which affects the sensitivity of the bandwidth allocated with respect to price changes. In the linear case, the cost component

is given by $\mathcal{C}'_i(\phi_i) = b_i D_i$, whereas an example of a nonlinear pricing function is $\mathcal{C}'_i(\phi_i) = b_i^{(e^{D_i} - d_i)} D_i$. In the latter case, the penalty cost increases exponentially when the desired delay cannot be met, compared with the linear pricing model.

Similarly, the type of service level requirements used has a significant impact on the amount of allocated bandwidth. Notice that the quality of service constraints can have either a deterministic form – e.g., average queue delay, $\bar{D}_i(\phi) < d_i$ – or a stochastic one – e.g., $P\{D_i(\phi) > d_i\} \leq \varepsilon_i$ [14,19]. It can also be seen that the constraints we allow in our formulation could correspond to any statistic of interest of the delay (or loss) distribution, such as the mean, the median or any other quantile of interest.

4. MBORA algorithm based on linear pricing and average queue delay guarantees

In this section, we investigate the optimal solution for the profit center under a linear pricing model and service requirement constraints expressed in terms of average queue delays.

Furthermore, we assume that the applications under of the generalized framework can be grouped according to their utility functions into the following four classes: *hard real-time*, *delay-adaptive*, *rate-adaptive* and *elastic*, given the generalized service model proposed by Shenker [13]. This generalized service model gives a well-formed abstraction of service model for Internet, so that our proposed framework under this generalized service model can be implemented under various differentiated network schemes, such as IntServ or DiffServ. Due to the deterministic delay-bound requirement of the *hard real-time* class, this class is not treated differently within the context of the MBORA system. Mapping classes accordingly into Fig. 1, we study the scenario with $m = 1$ and $n = 3$. Some remarks about solving the general problem of n classes, with the methodology developed in subsequent sections, are given in Section 4.3.

Under these assumptions, the provider's profit problem in this scenario is given by

$$\begin{aligned} & \max_{\phi} \left\{ \sum R_i - \sum \mathcal{C}'_i \right\} \\ & = \max_{\phi} \left\{ \sum_{i=1}^3 p_i \phi_i C' - \sum_{i=1}^2 \frac{b_i \bar{q}_i}{\phi_i C'} \right\}. \end{aligned} \quad (2)$$

subject to

$$\begin{aligned} & \sum_{i=1}^3 \phi_i \leq 1, \quad \phi_i \leq 1 \\ & \phi_i \geq \max\left\{\frac{r_i}{C'}, \frac{\bar{q}_i}{C' d_i}\right\}, \quad i = 1, 2, \quad \phi_3 \geq \frac{r_3}{C'} \end{aligned}$$

where \bar{q}_i is the average queue length, p_i is the price per unit of the utilization of the system's resources for class i and b_i is the cost per unit of time incurred by class i , and with $\phi = (\phi_1, \phi_2, \phi_3)$. Since the higher priority class requests a better service level than the lower priority classes, it should be charged with a higher price, while at the same time the provider has to reimburse the users at a higher rate for constraint violations; therefore, $p_i > p_j$ and $b_i > b_j$ for $i < j$. Note also that the delay cost for the elastic class is not considered, since the corresponding users are deemed insensitive to delays.

The quantity r_i that appears in the constraints is an estimated generalized traffic rate for class i . It could be selected as the mean rate, peak rate, or *effective bandwidth* (of one of many possible definitions), depending on the different service level requirements for different classes and how

conservatively the capacity is allocated among them for the shared-link system. The effective bandwidth is defined by Kelly as the bandwidth required to satisfy service constraints via large deviation theory [5]; that is a value between the mean rate and the peak rate. It is an efficient measure to allocate bandwidth in order to make their loss probability under a given bound [7]. For those classes without strict service level constraints, like our “elastic” class, the mean rate is adequate for bandwidth allocation purposes, whereas the effective bandwidth rate might be too generous.

The average queue length plays an important role in the above formulation and is derived from the fluid model given in [10]. Notice that the instantaneous queue length process at time t for class i can be obtained through the formula

$$q_i(t) = \max [q_i^0 + (r_i - \phi_i C')t, 0] \tag{3}$$

where q_i^0 is the initial queue length of the class i and t denotes the length of the time interval. In our proposed scheme, the share of system resources (bandwidth) allocated to the various classes would be dynamically assigned over an adaptive window W . Thus, the average queue length of class i during an adaptive window W is given by

$$\bar{q}_i = \frac{1}{W} \int_0^{\tau_i} q_i(t)dt = \frac{\tau_i}{W} \left[q_i^0 + \frac{\tau_i}{2} (r_i - \phi_i^d C') \right] \tag{4}$$

where $[0, \tau_i]$ is the time interval during which the queue length process remains positive during an adaptive window W and τ_i is determined by $\tau_i = \min\{t_i^0, W\}$ with t_i^0 being the time it takes to empty the queue. In turn, t_i^0 could be obtained by $t_i^0 = q_i^0 / (\phi_i C' - r_i)$, given the initial queue length q_i^0 . It can be seen that the average queue length depends on τ_i and therefore we distinguish the following cases: $t_i^0 < W$ and $t_i^0 > W$.

In the first case, the service constraint leads to the relationship

$$\phi_i > \frac{q_i^0}{WC'} + \frac{r_i}{C'} = \phi_i^1. \tag{5}$$

On the other hand, the average queue length \bar{q}_i for this case can be written as:

$$\bar{q}_i = \frac{[q_i^0]^2}{2W} \times \frac{1}{\phi_i C' - r_i}$$

Then, the delay component of the service constraint leads to the relationship

$$\phi_i \geq \frac{r_i}{2C'} + \sqrt{\left(\frac{r_i}{2C'}\right)^2 + \left(\frac{q_i^0}{C'}\right)^2} \times \frac{1}{2d_i W} = \phi_i^2. \tag{6}$$

Hence, equations 5 and 6, imply that $\phi_i \geq \max\{\phi_i^1, \phi_i^2\}$, while the constraint $\phi_i \leq 1$ has to be satisfied.

In the second case, we analogously obtain

$$\phi_i \leq \frac{q_i^0}{WC} + \frac{r_i}{C} = \phi_i^1 \tag{7}$$

In addition,

$$\bar{q}_i = q_i^0 + \frac{W}{2}(r_i - \phi_i C)$$

which in turn gives that

$$\phi_i \geq \frac{\frac{q_i^0}{WC} + \frac{r_i}{2C}}{\frac{1}{2} + \frac{d_i}{W}} = \phi_i^d \tag{8}$$

However, in some cases ϕ_i^d may be greater than ϕ_i^1 , which creates a conflict between the constraints. In such a case we must at least have $\phi_i \geq \frac{r_i}{C}$, whereas if ϕ_i^d is no greater than ϕ_i^1 , then the lower bound of ϕ_i should be given by the maximum of $\frac{r_i}{C}$ and ϕ_i^d .

Hence, an analogous reasoning leads to the following constraint $\phi_i^3 \leq \phi_i \leq \phi_i^1$, where

$$\phi_i^3 \begin{cases} \frac{r_i}{C} & \text{if } \phi_i^d > \phi_i^1 \\ \max[\frac{r_i}{C}, \phi_i^d] & \text{if } \phi_i^d \leq \phi_i^1 \end{cases}$$

4.1. Optimal allocation of resources

We start by considering two cases, specifically, whether the estimated aggregate traffic rate over all classes exceeds the excess bandwidth, C' , or not. In Section 4.2, a fast algorithm is presented that solves the problem at hand. We examine next the two cases in detail:

4.1.1. The over-provisioned case

If $E \sum r_i < C'$, that means that the excess bandwidth of the network element is enough to accommodate all the incoming jobs from the various classes during the next adaptive window, W . Equation 2 shows that the problem under consideration is a nonlinear optimization one with inequality constraints, that can be solved by Kuhn-Tucker conditions.

However, we explore the structure of the problem to identify the solution in a faster and more economical fashion, which in addition provides insight into the nature of the allocation scheme. Some algebra reduces the problem to one involving only two decision variables, namely ϕ_1 and ϕ_2 , with the optimal solution satisfying the feasibility constraint $\phi_1 + \phi_2 = 1 - \frac{r_3}{C'}$. Notice that since the elastic class is not delayed constrained, it suffices to provide it with the minimally required bandwidth that satisfies the feasibility constraint; hence, $\phi_3 = \frac{r_3}{C'}$. Moreover, the relationship between the duration of the measurement window W and the duration of emptying the queue leads to four separate cases for the optimization problem at hand that can be examined individually. We elaborate next on how one of the cases is obtained. The remaining three can be obtained following an analogous reasoning and similar calculations. The detailed mathematical derivations of the remaining three cases are provided in [18].

Suppose that both decision variables (i.e., ϕ_1 and ϕ_2) satisfy $\phi_i \geq \max\{\phi_i^1, \phi_i^2\}$. The optimization problem can then be written as:

$$\max_{\phi_1, \phi_2} f(\phi_1, \phi_2) = p_1 \phi_1 C' + p_2 \phi_2 C' - h_1(\phi_1, \phi_2),$$

where

$$h_1(\phi_1, \phi_2) = -\frac{b_1 \times [q_1^0]^2}{2W(\phi_1 C' - r_1)\phi_1 C'} - \frac{b_2 \times [q_2^0]^2}{2W(\phi_2 C' - r_2)\phi_2 C'}$$

subject to the constraints

$$\begin{aligned} \sum_{i=1}^2 \phi_i &= 1 - \frac{r_3}{C'} \\ \phi_1 &\geq \max\{\varphi_1^1, \varphi_1^2\}, \quad \phi_2 \geq \max\{\varphi_2^1, \varphi_2^2\}. \end{aligned}$$

The remaining three cases are determined by the following three sets of constraints and characterized by different functional forms of the $h_j(\phi_1, \phi_2)$, $j = 2, 3, 4$ cost function. In the second case the sets of constraints is given by: $\sum_{i=1}^2 \phi_i = 1 - \frac{r_3}{C'}$, $\phi_1 \geq \max\{\varphi_1^1, \varphi_1^2\}$ and $\varphi_2^3 \leq \phi_2 \leq \varphi_2^1$, in the third case by $\sum_{i=1}^2 \phi_i = 1 - \frac{r_3}{C'}$, $\varphi_1^3 \leq \phi_1 \leq \varphi_1^1$ and $\phi_2 \leq \max\{\varphi_2^1, \varphi_2^2\}$ and finally in the fourth case by $\sum_{i=1}^2 \phi_i = 1 - \frac{r_3}{C'}$, $\varphi_1^3 \leq \phi_1 \leq \varphi_1^1$ and $\varphi_2^3 \leq \phi_2 \leq \varphi_2^1$.

4.1.2. The under-provisioned case

If $E \sum r_i > C'$, the excess bandwidth of the link is not enough to satisfy all classes for the next adaptive window. Also, our constrained optimization problem is not feasible due to the conflicts between the feasibility constraints and the service guarantees. In this case, a relaxation rule for the constraints is proposed and incorporated in the bandwidth allocation scheme.

The relaxation rule operates in two stages: at small scales, where relaxation of the constraints within a class occurs and at large scale, where relaxation between classes takes place.

For the relaxation within a class, the relationship between the constraints for the i -th class is as follows:

$$\max\{\varphi_i^1, \varphi_i^2\} \geq \min\{\varphi_i^1, \varphi_i^2\} \geq \varphi_i^3 \geq \frac{r_i}{C'}$$

Our proposal is to relax the constraints recursively in the above descending order, that also reflects the degrading of QoS guarantees inside a class through our discussion for these constraints before. Therefore, the system can provide the best QoS service for each class in the small scale. Regarding the relaxation of the constraints among the classes, our rule implies that the constraints of the higher class should not be relaxed until all of the constraints of the lower classes in the above order are relaxed. In this way, the higher classes of traffic can be guaranteed with the best service that the system can provide.

Finally, in order to avoid starvation of a particular class (that would happen for the best effort class, if for example, $\sum_{i=1}^2 r_i \geq C'$), the network provider can implement pre-specified thresholds to discipline the allocation for each class, as studied in our previous work [17]. In such instances, the service requirements of the delay-adaptive class should be satisfied at the expense of the remaining two classes, provided that it does not exceed a pre-specified threshold θ_1 . Furthermore, the service level requirement of the rate adaptive class should also be satisfied, provided that it does not exceed a different pre-specified threshold θ_2 .

4.2. Calculating the optimal solution

In this section, we briefly outline how the four cases of the optimization problem under consideration can be solved in over-provisioned scenario. For the objective function corresponding to the four possible cases, it can be shown that the Hessian matrix of second partial derivatives is negative definite. For example, for the objective function in case 1 the Hessian is given by

$$\mathbf{H} = \begin{pmatrix} \frac{b_1[q_1^0 C']^2 [\frac{1}{(\phi_1 C' - r_1)^3} - \frac{1}{(\phi_1 C')^3}]}{-r_1 W} & 0 \\ 0 & \frac{b_2[q_1^2 C']^2 [\frac{1}{(\phi_2 C' - r_2)^3} - \frac{1}{(\phi_2 C')^3}]}{-r_2 W} \end{pmatrix}$$

which, under the feasibility constraints, is negative definite (since all its eigenvalues are negative). Therefore, it can be concluded that the objective function is jointly concave and hence possesses a unique maximum (maybe at a boundary point), which can be obtained by solving for the classical Kuhn-Tucker conditions.

However, by further exploring the structure of the problem at hand we can obtain the optimal solution in a more inexpensive and easy to implement manner. We illustrate the main steps of the proposed approach on the problem defined in case 1. The other optimization problems (cases 2–4) can be solved in an analogous manner (the details can be found in [17]). By solving the feasibility constraint $\phi_1 + \phi_2 = 1 - \frac{r_3}{C'}$ for ϕ_2 and substituting that value in the objective function we find a new objective function of a single variable given by

$$g(\phi_1) = p_1 \phi_1 C' + p_2 (C' - \phi_1 C' - r_3) - \frac{b_1 [1_1^0]^2}{2W(\phi_1 C' - r_1)\phi_1 C'} - \frac{b_2 [q_2^0]^2}{2W(C' - \phi_1 C' - r_2 - r_3)(C' - \phi_1 C' - r_3)}$$

It can easily be obtained that $g''(\phi_1) < 0$, which implies that $g(\phi_1)$ is a concave function. The derivation of the $g'(\phi_1)$ helps us determine the optimal solution, as follows. First denote the lower bound of the feasible region by B_L and the upper bound by B_U . The lower and upper bounds are such that they satisfy the QoS constraints (see equations 5–6) for both ϕ_1 and ϕ_2 (see also Fig. 3). If $g'(B_L) > 0$ and $g'(B_U) > 0$, then the optimal solution is given at the boundary by B_U , whereas if $g'(B_L) < 0$ and $g'(B_U) < 0$, then the optimal solution is given at the other boundary point B_L . Finally, if $g'(B_L) > 0$ and $g'(B_U) < 0$, then the optimal solution lies in the interior of the interval (B_L, B_U) and must be found by numerical root finding methods, such as the bisection method, or Newton’s method ([3]).

The globally optimal solution is then obtained by calculating first the optimal solution $\phi_1^*(k), k = 1, 2, 3, 4$ for the 4 cases and then keeping the maximum amongst the four. Combined with our under-provisioned solution, the algorithm for the optimal solution is given in pseudo-code form next.

Algorithm 1. Obtaining the optimal solution

```

if  $\sum_{i=1}^3 r_i < C'$  then
    Identify the  $N$  feasible cases by checking the
    underlying feasibility constraints
    while  $N > 0$  do
    
```

```

Obtain  $B_i^L, B_i^U$  for case  $i$ 
if  $g'(B_L^i) > 0$  and  $g'(B_U^i) \geq 0$  then
     $\phi_1^{i*} = B_U^i$ 
else if  $g'(B_L^i) \leq 0$  and  $g'(B_U^i) < 0$  then
     $\phi_1^{i*} = B_L^i$ 
else
    Use root finding method for solving  $g'(\phi_1) = 0$ 
    to obtain  $\phi_1^{i*}$ .
end if
Obtain the maximum for case  $i$ ,  $\max f^{i*}(\phi_1^{i*}, \phi_2^{i*})$ 
 $N = N - 1$ 
end while
 $\max f^*(\phi_1^{i*}, \phi_2^{i*}) = \max_{i \in N} f^{i*}(\phi_1^{i*}, \phi_2^{i*})$ 
else
if  $(\frac{r_1}{C'} + \frac{r_2}{C'}) < 1$  then
    Constraints relaxation in Section 4.1.2
else
    The threshold  $(\phi_1, \phi_2)$  in [17]
end if
end if
    
```

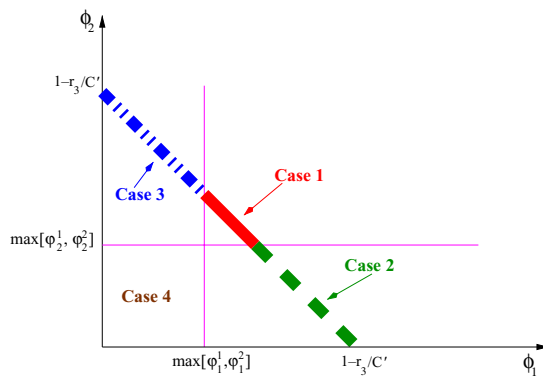
Remark: The importance of the above algorithm is that a complicated high-dimensional constrained nonlinear optimization problem is reduced to one involving a one-dimensional constrained concave function that can be tackled directly through standard calculus methods.

4.3. Complexity analysis and feasible regions

In this subsection, we will analyze the complexity of our proposed algorithm and discuss the feasible regions of the optimization problem.

Let us start with some insights about the feasible regions by examining the following plot. Suppose that the intersection of the constraints from case 1 occurs inside the region determined by the inequality $\phi_1 + \phi_2 \leq 1 - \frac{r_3}{C'}$ (see Fig. 3); we can then conclude that case 1 is feasible. It is also then easy to see that we do not have to consider the optimization problem given in case 4, due to the incompatibility between the constraints of those two cases. To summarize, the

Fig. 3 Structure of the overall optimization problem when case 1 (solid line) is feasible.



sufficient condition for case 1 is $\max\{\{\varphi_1^1, \varphi_1^2\} + \max\{\varphi_2^1, \varphi_2^2\}\} \leq 1 - \frac{r_3}{C}$. Analogous conclusions can be obtained by examining similar plots corresponding to the other possible constraints.

Therefore, in our proposed algorithm, there are at most three feasible cases for the over-provisioned scenario. In all, we can conclude that the number of feasible cases is determined by the interaction among the constraints for all classes for the over-provisioned case.

Next, we will discuss the computation complexity of our proposed algorithm. Suppose we extend our study to more generalized n classes. For the over-provisioned case, there would be 2^n cases for the objective functions. Similarly, we can prove the joint concavity for all cases, and its optimal solution can be obtained by Newton's method. It is obvious that the number of loops in our algorithm is relied on the number of feasible cases, that in turn can lead to the computation complexity of the proposed algorithm to be $O(2^n)$ in the *worst case*. However, the relationship between the service guarantee and the feasibility constraints reduces the number of cases to be examined for locating the optimal value considerably. With respect to the under-provisioned case, the computation complexity of our proposed algorithm is $O(2^n)$ because the relaxation procedure has to go through at most n classes, inside of which is just simple 'if-else' decision. However, it should be noticed that, in realistic circumstances, n is not a particularly large number (usually no larger than 10) finite number no larger than 10, considering the limited number of the aggregated traffic classes in networks. That makes the computation complexity of our proposed algorithm trivial, favored for online implementation. Finally, the problem of determining the feasible cases is a topic of current research.

4.4. Performance evaluation and comparisons

The model used in our experimental investigations is a processor sharing system with three input queues that map to the delay-adaptive, the rate-adaptive and elastic class, respectively. The link capacity is set at 1 Mb/s and the model parameters employed in the simulation are given in Table 1.

For the adaptive algorithms proposed in this paper, the measurement window was set to 1 sec. An on-off input Fractal Modulated Poisson process (FMPP) model, proposed in [1], was employed. Finally, for simplicity purposes, the packet size was fixed to 50 bytes and the Hurst parameter of the traffic process to 0.75. Several other simulation scenarios were considered in [17] that produced analogous results.

Given the link speed, the system capacity (stability) constraint is at 0.4 ms. That is, if the mean packet inter-arrival time is larger than this value, the system is over-provisioned and relatively few packet losses are expected to occur. On the other hand, for mean packet inter-arrival times smaller than 0.4 ms, the system is unstable and highly backlogged queues together with frequent packet losses are expected. The case where the mean time is exactly 0.4 corresponds to the critical regime.

In the ensuing discussion, the classical WRR and WFQ with static allocated weights are denoted by SWRR and SWFQ, respectively, whereas their dynamic counterparts are denoted by

Table 1 Parameters for Different Classes

	Delay-Adaptive	Rate-Adaptive	Elastic Class
Initial Weight	2	2	1
Buffer Size (bytes)	500	5000	10000
p (cents/kbps)	5	1	–
b (cents/ms)	0.5	0.1	–

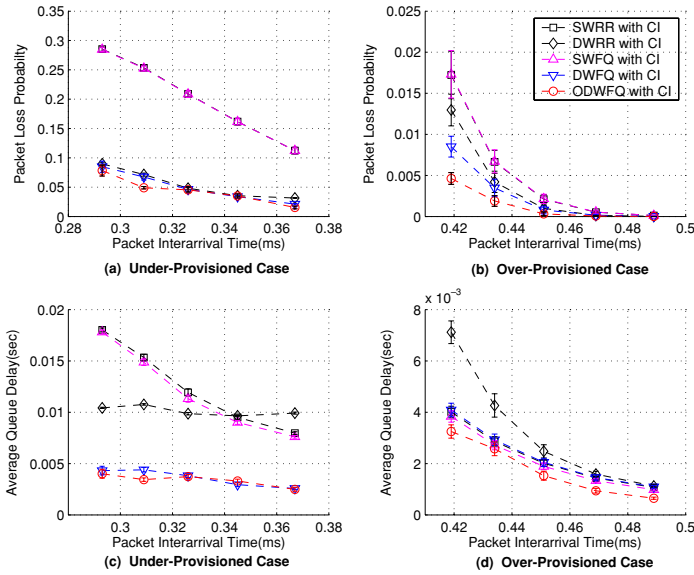


Fig. 4 Comparison of scheduling algorithms on packet loss probability in (a) under-provisioned and (b) over-provisioned cases; on average queue delay in (c) under-provisioned and (d) over-provisioned cases for the delay-adaptive class.

DWRQ (Dynamic Weighted Round Robin) and DWFQ (Dynamic Weighted Fair Queue). For the dynamic policies the weights are sequentially updated at the beginning of every W -window.

Finally, the profit based optimal policy presented in Section IV is denoted by DWFQ (Optimized Dynamic Weighted Fair Queue policy). The performance metrics used in our study are packet loss probabilities and average queue delays. In order to facilitate the presentation of the results we separate the under- from the over-provisioned scenarios, due to the incorporation of ‘non-starvation’ thresholds in the proposed policies. In the plots that follow, the class average loss probability and average delay together with 95% confidence intervals obtained from 50 replications are shown.

In the four panels of Fig. 4, the performance of all 5 algorithms under study for the delay-adaptive class is shown. Specifically, Figs 4 (a) and (b) show that the adaptive versions of the WRR and WFQ outperform with respect to losses their static counterparts, for both the under- and over-provisioned scenarios. The ODWFQ exhibits the best performance in all cases, and in particular for stable but heavily loaded systems (range between 0.42 and 0.44 ms). Finally, as the system becomes more lightly loaded (0.48 ms and beyond) all policies become essentially equally efficient. Regarding the average delay metric (Figs 4 (c) and (d)), the ODWFQ dominates all other policies in the stable case and performs very well in the under-provisioned scenario. From the remaining policies, we notice the strong performance of the DWFQ policy, and the inferior performance of the DWRR policy. The latter finding is due to the inaccuracy involved with the normalization of the weights.

In the panels of Fig. 5, the performance of the rate-adaptive class is examined. For this class the ODWFQ policy clearly outperforms all its competitors, for both performance metrics in the under- and over-provisioned cases. Among the remaining policies the DWFQ outperforms its static counterpart in all cases. On the other hand, the pattern for the round robin policies is more involved. For example, the DWRR clearly outperforms the SWRR policy for both performance

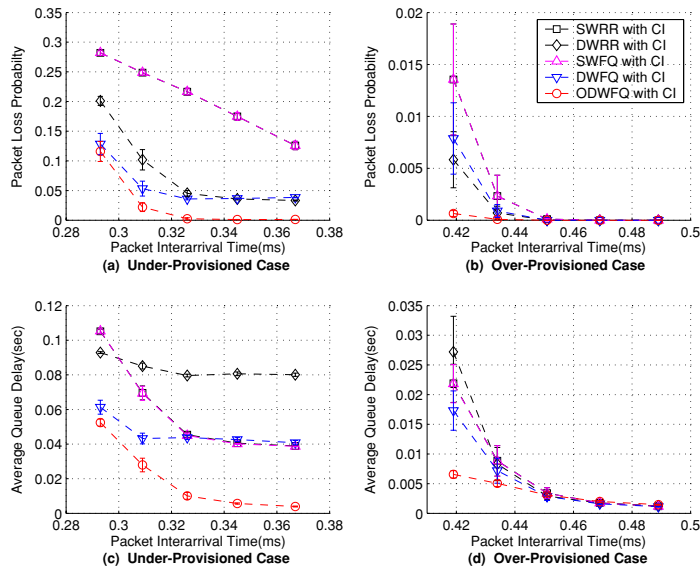


Fig. 5 Comparison of scheduling algorithms on packet loss probability in (a) under-provisioned and (b) over-provisioned cases; on average queue delay in (c) under-provisioned and (d) over-provisioned cases for the rate-adaptive class.

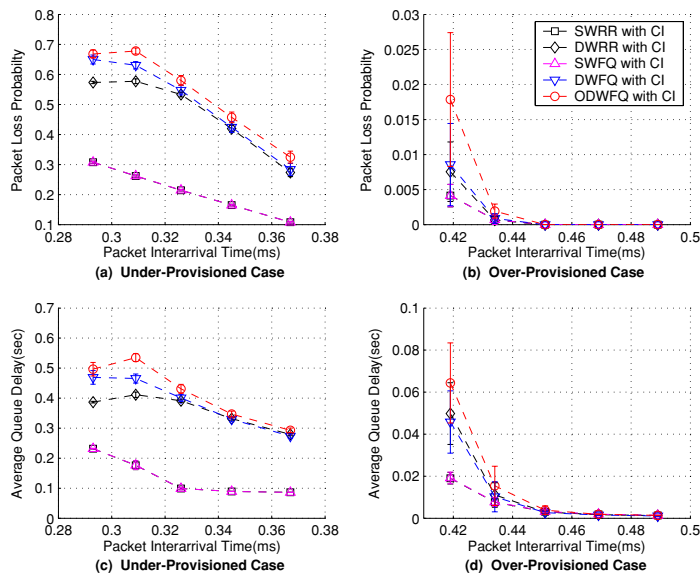


Fig. 6 Comparison of scheduling algorithms on packet loss probability in (a) under-provisioned and (b) over-provisioned cases; on average queue delay in (c) under-provisioned and (d) over-provisioned cases for the elastic class.

metrics in the under-provisioned case and for the loss metric in the over-provisioned case; however, the SWRR policy exhibits a better performance with respect to delay for the over-provisioned case. Finally, it should be noted that as the incoming traffic's rate decreases (the mean interarrival time increases) and the system's load becomes lighter, the performance of all three policies become fairly similar.

In the panels of Fig. 6 the results for the elastic class are shown. For this class the ODWFQ exhibits the worst performance for all cases and metrics, due to the fact that this class yields very low profits and hence less bandwidth is allocated to it under this policy. For this class, the static policies achieve the best performance in terms of loss and average queue delay for the elastic class in both under-provisioned and over-provisioned cases, since the dynamic policies allocate less bandwidth to this class, in order to meet the service requirements of the higher priority classes.

In conclusion, we observe that the proposed ODWFQ policy exhibits the best and most robust performance for both the over- and under-provisioned scenarios for the high paying traffic classes.

5. Conclusions

Adaptive scheduling based on measurements of traffic and the system's queueing state has the potential of greatly improving the efficiency of resource allocation techniques in emerging flexible networks. In our previous work on this topic, we have introduced a measurement-based adaptive scheduler and validated its performance with extensive simulation results. In this paper, we have formulated the generalized online setting of adaptive schedulers as a formal optimization problem taking into account service constraints and the underlying pricing scenario. Based on the simplest LP+AQD model, we then proceeded to study its solutions on a case-by-case basis, thus obtaining fundamental insights regarding the implementation and use of such schemes.

Acknowledgements The authors would like to thank the special issue editor Harry Perros and the anonymous referee for their useful comments that improved the presentation of the material. The work of George Michailidis has been supported in part by NSF through grants IIS-9988095 and CCR-0325571. The work of Peng Xu and Michael Devetsikiotis was supported through a grant from Alcatel, Research and Innovation, Piano.

References

1. B. Ryu and S. Lowen, Fractal Traffic Models for Internet Simulation, in: *IEEE International Symposium on Computers Communications*, France (2000) 271–370.
2. J. C.R. Bennett and H. Zhang, Hierarchical Packet Fair Queueing Algorithms, *IEEE/ACM Trans. Networking* (Oct 1997) 675–689.
3. R.L. Burden and J.D. Faires, *Numerical Analysis*, Brooks-Cole Publishing, seventh edition, 2001.
4. A. Chandra, W. Gong and P. Shenoy, Dynamic Resource Allocation for Shared Data Centers Using Online Measurements, in: *Proceedings of ACM/IEEE Intl Workshop on Quality of Service (IWQoS)* (2003), 381–400.
5. F. Kelly, Notes on Effective Bandwidth, in: F. P. Kelly, S. Zachary and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, Oxford University Press (1996), pp. 141–168.
6. L. Georgiadis, R. Guerin, V. Peris and K. N. Sivarajan, Efficient Network QoS Provisioning based on Per Node Traffic Shaping *IEEE/ACM Trans. Networking* 4(4) (1996) 482–501.
7. R. Guerin, H. Ahmadi and M. Naghshineh, Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks, *IEEE J. Select. Areas Commun* 9(7) (Sept. 1991) 968–981.
8. E. Knightly and N. Shroff, Admission Control for Statistical QoS: Theory and Practice, *IEEE Network* 13(2) (1999) 20–29.
9. L. Breslau, S. Jamin and S. Shenker, Comments on the Performance of Measurement-based Admission Control Algorithms, in: *Proc. of IEEE INFOCOM* (2000) pp. 1233–1242.

10. Y. Liu, F.L. Presti, V. Misra, D. Towsley and Y. Gu, Fluid Models and Solutions for Large-Scale IP Networks, in: *Proceedings of ACM/SIGMETRIC 2003* (2003).
11. A. Parekh and R. Gallager, A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case, *IEEE/ACM Trans. Networking* (June 1993) 344–357.
12. J. Qiu and E. Knightly, Measurement-based Admission Control with Aggregate Traffic Envelopes.
13. S. Shenker, Fundamental Design Issues for the Future Internet, *IEEE Journal on Selected Areas in Communication*. 13(7) (1995) 1176–1188.
14. V. Sivaraman and F.M. Chiussi, Providing End-to-End Statistical Delay Guarantees with Earliest Deadline First Scheduling and Per-Hop Traffic Shaping, in: *Proc. of IEEE INFOCOM 2000* (2000).
15. J. Walrand, *An Introduction to Queuing Networks*. Prentice Hall, 1988.
16. H. Wang, C. Shen and K.G. Shin, Adaptive-Weighted Packet Scheduling for Premium Service, in: *Proceedings of the IEEE ICC*, 2001.
17. P. Xu, M. Devetsikiotis and G. Michailidis, Adaptive Scheduling using Online Measurements for Efficient Delivery of Quality of Service. Technical Report TR2004-12, SAMSI, RTP, NC, March 2004.
18. P. Xu, G. Michailidis and M. Devetsikiotis, Online Scheduling for Resource Allocation of Differentiated Services: Optimal Settings and Sensitivity Analysis. Technical Report TR2004-13, SAMSI, RTP, NC, March 2004.
19. Z.L. Zhang and J.K.D. Towsley, Statistical Analysis of Generalized Processor Sharing Scheduling Discipline, *IEEE J. Select. Areas Commun.* 13(6) (1995) 1071–1080.