Optimal Robust Policies for Bandwidth Allocation and Admission Control in Wireless Networks

Vicent Pla[*]
Univ. Politécnica de Valencia
ETSIT, Camí de Vera s/n
46022 Valencia, Spain
vpla@dcom.upv.es

Jorma Virtamo
Dept. of Communications and Networking
Helsinki Univ. of Technology
P.O. Box 3000
FI02015 TKK, Finland
jorma.virtamo@tkk.fi

Jorge Martínez-Bauset[†]
Univ. Politécnica de Valencia
ETSIT, Camí de Vera s/n
46022 Valencia, Spain
jmartinez@upvnet.upv.es

ABSTRACT

We consider joint strategies of bandwidth allocation and admission control for elastic users competing for a downlink data channel in a cellular network. For the sake of robustness and generality of the results we focus on the set of strategies whose performance does not depend on the detailed traffic characteristics beyond the traffic intensity. Performance is studied at the flow level in a dynamic setting where users come and go over time. A number of user classes are considered, which are characterized by their achievable bit rate, guaranteed throughput, arrival rate and mean flow size. We aim at characterizing a strategy which is optimal in the sense of having the lowest blocking probability. Such characterization provides some interesting insights into the optimal policy and its evolution as the system load increases. Unfortunately, from a practical perspective computing the optimal policy can be exceedingly complex except for lightly loaded systems. Alternatively, we propose a computationally feasible suboptimal policy that achieves a good relative performance.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Queueing theory, Stochastic processes; G.1.6 [**Numerical Analysis**]: Optimization; C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication*; C.2.3 [**Computer-Communication Networks**]: Network Operations—*Network management*

General Terms

Performance

Keywords

Wireless networks, resource management, elastic traffic, insensitivity, optimal policy

1. INTRODUCTION

In spite of the enormous variety of traffic flows in the multiservice Internet and the corresponding difficulty in its characterization, the abstraction of flows pertaining to two broad classes (*elastic* and *streaming*) has proven to be simple yet practical for traffic engineering purposes [26]. Future wireless cellular networks are expected to provide not only voice service but also data services —mainly Internet access traffic— thus carrying the same traffic type than wired access networks, although not necessarily in the same proportion.

Streaming traffic corresponds to real-time audio or video and hence has rather stringent requirements on packet delay and jitter. Elastic traffic in turn can adapt to available bandwidth up to a minimum. If available bandwidth drops below that minimum flows may abandon before completing the transaction leading to an unwanted waste of resources [23, 12]. Therefore, despite the adaptability of elastic flows it is advisable to enforce some type of admission control (AC) in order to guarantee a minimum bandwidth per flow and ensure an efficient use of resources.

In the quest of an equivalent of the Erlang formula for the Internet, Bonald and Proutiére, following the seminal work of Kelly [20], invented the concept of *balanced fairness* [9, 10], which is defined as the most efficient way to share network capacity among different flows so that (under some assumptions on the generation of flows) the resulting system is insensitive, i.e., the stationary distribution of the underlying queuing network does not depend on the detailed distribution of service requirements beyond their means. A further extension of the results on insensitive queuing network is presented in [4]. In [8] balanced fairness allocation is applied to several network types and its performance is compared with that of *max-min fairness* and *proportional fairness*.

Despite the packetized nature of data in 3G networks, performance analysis of wireless data networks at the flow level has attracted a considerable interest during the last

[*] This author was supported by the Spanish Government (30% PGE) and the European Commission (70% FEDER) through project TSI2007-66869-C02-02 and by the European Commission through NoE EuroNGI.

[†] This author was supported by the Spanish Government (30% PGE) and the European Commission (70% FEDER) through project TSI2005-07520-C03-03.

ValueTools 2008, October 21 – 23, 2008, Athens, GREECE.
Copyright ©2008 ICST ISBN # 978-963-9799-31-8.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376

years [5, 6, 11, 14, 17, 15, 16, 22, 28]. The vast majority of these studies rely to some extent on queuing models based on Processor Sharing (PS) and Discriminatory Processor Sharing (DPS) queues; see, for instance, the seminal paper [18] and [1, 2, 3, 29].

In [14, 17] it is shown that the *Proportional Fair* scheduler, which is commonly deployed in downlink data channels, at the flow level can be satisfactorily evaluated by means of a Processor Sharing (PS) based queuing model. The resulting model has the advantage of being insensitive. Applications of such model to a single cell with no moving users can be found in [11, 14, 17].

On the other hand, DPS based models are introduced to model the unequal capacity sharing arising in some situations. Unequal sharing may arise fundamentally due to TCP rate control or service differentiation enforced by packet schedulers. However, in normal load conditions, realizing service differentiation through a packet scheduler that operates in a DPS manner is rather ineffective [13].

In this paper we focus on elastic traffic carried over a wireless cellular network. Specifically, we address jointly the problem of bandwidth allocation (BWA) and admission control (AC). Our main goal is to characterize the optimal joint BWA-AC scheme —in the sense of having the lowest loss probability— among those that are insensitive to the distribution of the flow size, i.e., only depends on their mean values.

Our work inherits some of the ideas of a series of papers dealing with insensitive dynamic load balancing [7, 19, 21]. In all of them —and also here— the simplicity and robustness of insensitivity is an essential condition for the optimal policy that is sought. In [7] it is assumed that capacities are allocated according to balanced fairness and then the optimal routing policy is sought constrained to being balanced in order to preserve insensitivity. The optimality objective is to minimize the overall blocking probability or the maximum per-class blocking probability. A simple characterization of the optimal routing policy is obtained for the single-class traffic and also for the more general multi-class traffic. However, in the latter case the policy optimization is restricted to the set of *decentralized policies*, i.e., strategies where the routing decision for a class-k customer does not depend on the number of customers of other classes.

For the purpose of obtaining the insensitivity property it is not necessary that capacity allocation and routing are balanced separately. Actually, it was already noted in [9] that a better performance can be achieved if capacity allocation and routing are jointly balanced, which is a weaker requirement than separate balancing. This approach is followed in [19] and [21]. However the performance advantage of joint balancing comes at the cost of higher complexity. In [19] the authors obtain and characterize the optimal joint allocation-routing policy in a single-class traffic scenario. To best of our knowledge no similar results exists for the multiclass traffic scenario. A multiclass traffic scenario with global information policies is studied in [21] but the aim is not a characterization of the optimal policy. An approach based on the theory of Markov Decision Processes (MDP) is used to formulate the optimization as a Linear Programming (LP) problem (see, for instance, [27]). The LP formulation allows more flexibility and hence a greater variety of problems, objective functions and constraints can be considered [21].

In this paper we address the more general problem of seeking a characterization of the optimal global policy in a multiclass traffic setting, but in turn we restrict ourselves to a simpler network topology — which fits a cellular scenario — than in all the aforementioned studies of this kind [7, 19, 21]: all traffic classes have a single feasible route into which they are allocated or otherwise blocked, and there is a single constraining resource, i.e. the time-slotted wireless channel. We employ the same optimization approach as in [21] based on an MDP-LP formulation.

The remainder of the paper is structured as follows. Section 2 describes the model of the system and the optimization problem is formally stated. In Section 3 we present the characterization of the optimal policy and introduce a suboptimal policy. Numerical experiments illustrating the contents of this section are shown in Section 4. Concluding remarks are given in Section 5.

2. MODEL DESCRIPTION AND PROBLEM FORMULATION

We model traffic at the flow level and ignore interactions at the packet level (scheduling, buffer management, TCP congestion control,...). The flow content is then viewed as a fluid which is transmitted as a continuous stream with rate changes occurring only at flow arrivals and departures. This is a widely used traffic model in the literature (see for instance [8] and its references). We focus on a single base station with a downlink channel allocated to data users. We consider that the downlink resources are time-shared among active users, i.e., flows. Transmission is done in a one-by-one fashion using time slots with duration much shorter than flow duration or flow inter-arrival times so that the validity of the flow level abstraction is maintained [11, 28].

Flows are classified into K different classes. Class-i flows arrive as a Poisson process with rate λ_i, their mean flow size (expressed in bits) is $1/\mu_i$ and require a minimum bit rate φ_i. Let C_i denote the feasible bit rate for class-i flows, i.e., the bit rate that is achieved during a slot assigned to one of such flows. Moreover, let us introduce $\rho_i = \lambda_i/(C_i\mu_i)$, $\rho = \sum_{i=1}^{K} \rho_i$ and $\lambda = \sum_{i=1}^{K} \lambda_i$. Classes can be defined by the feasible rates —which correspond to different locations within the cell [11, 25]—, flow types —having different mean flow sizes or minimum rate requirements—, or both.

Let $\boldsymbol{x} = (x_1, \ldots, x_K)$ denote the system state, where x_i is the number of active flows of the i-th class. The BWA-AC policy is described by $\phi_i(\boldsymbol{x})$ and $p_i(\boldsymbol{x})$: when the system is at state $\boldsymbol{x}$ arriving class-i flows are accepted with probability $p_i(\boldsymbol{x})$ and the ensemble of class-i flows is served with bit rate $\phi_i(\boldsymbol{x}) = C_i\tau_i(\boldsymbol{x})$, i.e., a fraction $\tau_i(\boldsymbol{x})$ of time-slots is assigned to class i; note that

$$\sum_{i=1}^{K} \tau_i(\boldsymbol{x}) = 1. \tag{1}$$

Within a class, bandwidth is fairly shared among the flows. The bit rate seen by a class-i flow is $\phi_i(\boldsymbol{x})/x_i$ (if $x_i > 0$).

Subject to the minimum bit rate per flow requirements $\phi_i(\boldsymbol{x}) \geq x_i\varphi_i$ and the total capacity constraint (1), it is easily seen that the set of feasible states is

$$\mathcal{S} := \left\{ \boldsymbol{x} : \quad \sum_{i=1}^{K} \frac{x_i}{\beta_i} \leq 1 \right\},$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376

where $\beta_i = C_i/\varphi_i$.

Denote by $\pi(\boldsymbol{x})$ the stationary state probabilities and by P_b the aggregate blocking probability. We want to find the insensitive BWA-AC policy that minimizes P_b while fulfilling the minimum rate requirements. More formally the problem can be stated as:

Find: $\phi_i(\boldsymbol{x})$ and $p_i(\boldsymbol{x})$ for $i = 1, \dots, K$ and $\boldsymbol{x} \in \mathcal{S}$ that

Minimize: P_b

Subject to:

1. insensitivity with respect to the flow size distribution;

2. minimum rate requirements: $\phi_i(\boldsymbol{x}) \geq x_i \varphi_i$, $\forall i$.

We formulate the optimization problem above as an MDP-LP. The state of the MDP consists of the system state $\boldsymbol{x}$, the admission decision $\boldsymbol{d}$ vector, and the bandwidth allocation b variable. The admission vector $\boldsymbol{d} = (d_1, \dots, d_K) \in \{0,1\}^K$ codes which traffic classes will have their newly arriving flows accepted: if $d_i = 1$ new class-i flows are accepted, and rejected otherwise. The bandwidth allocation variable codes to which class the transmission capacity is allocated: $b = i$ means that transmission capacity is allocated to class i. Let $\pi(\boldsymbol{x}, \boldsymbol{d}, b)$ denote the MDP state probability, in other words, the probability that the system is at state $\boldsymbol{x}$, accepts only those new flows belonging to classes in the set $\{i : d_i = 1\}$, and the transmission capacity is allocated to ongoing class-b flows.

The system state probabilities $\pi(\boldsymbol{x})$, blocking probability P_b and policy parameters $\phi_i(\boldsymbol{x})$ and $p_i(\boldsymbol{x})$ can be expressed in terms of $\pi(\boldsymbol{x}, \boldsymbol{d}, b)$ as

$$\pi(\boldsymbol{x}) = \sum_{\boldsymbol{d} \in \{0,1\}^K} \sum_{b=1}^{K} \pi(\boldsymbol{x}, \boldsymbol{d}, b),$$

$$P_b = \sum_{i=1}^{K} \left(\frac{\lambda_i}{\lambda} \sum_{\boldsymbol{d}: d_i = 0} \sum_{\boldsymbol{x} \in \mathcal{S}} \sum_{b=1}^{K} \pi(\boldsymbol{x}, \boldsymbol{d}, b) \right),$$

$$\tau_i(\boldsymbol{x}) = \frac{\sum_{\boldsymbol{d}} \pi(\boldsymbol{x}, \boldsymbol{d}, i)}{\sum_{\boldsymbol{d}} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b)}, \tag{2}$$

$$p_i(\boldsymbol{x}) = \frac{\sum_{\boldsymbol{d}: d_i = 1} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b)}{\sum_{\boldsymbol{d}} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b)}. \tag{3}$$

The LP problem can now be written as follows

$$\min_{\pi(\boldsymbol{x}, \boldsymbol{d}, b)} \sum_i \left(\frac{\lambda_i}{\lambda} \sum_{\boldsymbol{d}: d_i = 0} \sum_{\boldsymbol{x}} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b) \right), \tag{4}$$

subject to:

$$\pi(\boldsymbol{x}, \boldsymbol{d}, b) \geq 0 \qquad \forall \boldsymbol{x} \in \mathcal{S},\ \boldsymbol{d} \in \{0,1\}^K,\ b = 1, \dots, K, \tag{5}$$

$$\sum_{\boldsymbol{x}} \sum_{\boldsymbol{d}} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b) = 1, \tag{6}$$

$$\beta_i \sum_{\boldsymbol{d}} \pi(\boldsymbol{x}, \boldsymbol{d}, i) \geq x_i \sum_{\boldsymbol{d}} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b) \qquad \forall \boldsymbol{x} \in \mathcal{S},\ i = 1, \dots, K, \tag{7}$$

$$\rho_i \sum_{\boldsymbol{d}: d_i = 1} \sum_b \pi(\boldsymbol{x} - \boldsymbol{e}_i, \boldsymbol{d}, b) = \sum_{\boldsymbol{d}} \pi(\boldsymbol{x}, \boldsymbol{d}, i) \qquad \forall i = 1, \dots, K, \quad \boldsymbol{x} \in \mathcal{S} : x_i > 0, \tag{8}$$

where $\boldsymbol{e}_i$ is the vector with a 1 in the i-th position and 0's elsewhere.

Equations (5) and (6) refer to probabilistic nature of $\pi(\boldsymbol{x}, \boldsymbol{d}, b)$, Eq. (7) represents the minimum rate requirement and Eq. (8) is the detailed balance condition. In the ordinary LP formulation of MDP theory, global balance conditions appear as linear constraints on the decision variables. In order to retain insensitivity, we impose stricter detailed balance conditions as constraints [21], which is equivalent to the balance condition [21, 4]

$$\frac{\psi_i(\boldsymbol{x} - \boldsymbol{e}_j)}{\psi_i(\boldsymbol{x})} = \frac{\psi_j(\boldsymbol{x} - \boldsymbol{e}_i)}{\psi_j(\boldsymbol{x})} \qquad i, j = 1, \dots, K, \quad \boldsymbol{x} \in \mathcal{S} : x_i, x_j > 0,$$

where

$$\psi_i(\boldsymbol{x}) = \rho_i \frac{p_i(\boldsymbol{x} - \boldsymbol{e}_i)}{\tau_i(\boldsymbol{x})}. \tag{9}$$

Note that the radio channel capacity constraint is implicitly included in the definition of $\pi(\boldsymbol{x}, \boldsymbol{d}, b)$. From (2) it readily follows that $\sum_{i=1}^{K} \tau_i(\boldsymbol{x}) = 1$, actually it also holds for $\boldsymbol{x} = (0, \dots, 0)$, although it has no physical sense.

3. POLICY CHARACTERIZATION

For a given configuration, the LP formulated in the previous section can be numerically solved to obtain the values of $\pi(\boldsymbol{x}, \boldsymbol{d}, b)$ and by applying Eqs. (2)–(3), the BWA-AC parameters are obtained.

Our goal is to find a characterization for the optimal insensitive joint BWA-AC policy. In our quest we followed an inductive and rather experimental process: from the observation of particular solutions in rather simple scenarios we extracted and generalized the underlying characteristics of the optimal policy, which have been subsequently tested against a variety of more complex settings. In this section we describe the general form of the optimal insensitive joint BWA-AC policy. Since in some instances the general form may turn out to be excessively complicated for practical purposes, we also describe a simpler suboptimal form.

Denote by $\hat{\boldsymbol{\rho}} = (\hat{\rho}_1, \dots, \hat{\rho}_K) = \rho^{-1} \boldsymbol{\rho}$ the traffic share across classes. Let us denote by letter ω with a subscript a BWA-AC policy, i.e., a set of values for $\{\tau_i(\boldsymbol{x}), p_i(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}, i = 1, \dots, K\}$. Let $\omega(\rho)$ represent the optimal policy as a function of the system load ρ. It has been found that, for a given traffic share $\hat{\boldsymbol{\rho}}$, there exists a finite number of thresholds for ρ

$$0 = \rho^{(0)} < \rho^{(1)} < \rho^{(1)} < \cdots < \rho^{(m)} = \infty,$$

such that $\omega(\rho) = \omega_j$ for $\rho \in [\rho^{(j-1)}, \rho^{(j)}]$. Therefore, $\omega(\rho)$ (and thus $\tau_i(\boldsymbol{x})$ and $p_i(\boldsymbol{x})$) is a piecewise constant function of ρ. Moreover, as it will be seen below, the policy settings ω_j do not depend on the load conditions (ρ_i), they only depend on the values of C_i and φ_i. On the contrary the load thresholds $\rho^{(j)}$ do depend on the load conditions. In Section 3.1 we precisely specify the form of ω_1 and in Section 3.2 we describe the transformations that ω_1 undergoes as ρ increases giving rise to $\omega_2, \cdots, \omega_m$.

On the other hand, if the policy specification is available then the values of $\psi_i(\boldsymbol{x})$ can be easily computed (see Eq. (9)) and from these the system state probabilities easily follow as

$$\pi(\boldsymbol{x}) = \pi(\boldsymbol{0}) \psi_{i_1}(\boldsymbol{e}_{i_1}) \psi_{i_2}(\boldsymbol{e}_{i_1} + \boldsymbol{e}_{i_2}) \cdots \psi_{i_n}(\boldsymbol{x}). \tag{10}$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376

Where $\mathbf{0}=(0,\dots,0)$, $n=\sum_{i=1}^{K} x_i$ is the number of flows in the state $\boldsymbol{x}$, and

$$\langle \mathbf{0}, \boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_1}+\boldsymbol{e}_{i_2},\dots,\boldsymbol{e}_{i_1}+\dots+\boldsymbol{e}_{i_n}=\boldsymbol{x}\rangle ,$$

is any direct path from state $\mathbf{0}$ to state $\boldsymbol{x}$. Note that the simple product-form above for the system state probabilities is another consequence of the detailed balance condition. The blocking probability can be then computed by

$$P_b=\sum_{i=1}^{K}\left(\frac{\lambda_i}{\lambda}\sum_{\boldsymbol{x}\in\mathcal{S}}\left(1-p_i(\boldsymbol{x})\right)\pi(\boldsymbol{x})\right).$$

In principle having the piecewise characterization of $\omega(\rho)$ does not save having to solve the LP since the load thresholds $\rho^{(j)}$ remain to be known, but it can be circumvented and the exact policy to apply can be determined as

$$\omega(\rho)=\arg\min_{\omega_j} P_b(\omega_j,\rho). \tag{11}$$

This approach can be especially convenient if working with suboptimal policies (see Section 3.3 below).

3.1 The *First Policy* ω_1

For a sufficiently low load the optimal policy is ω_1, i.e., $\omega(\rho)=\omega_1$ if $0\le\rho<\rho^{(1)}$. Here we describe the observed principles that characterize ω_1 and by applying those principles we obtain a method for computing the policy parameters.

Throughout this subsection we assume, without loss of generality, that $C_1\ge C_2\ge\dots\ge C_K$. Define $\kappa(\boldsymbol{x})=\max\{i:x_i>0\}$.

At any state $\boldsymbol{x}$, the observed principles can be stated as:

1. The constraining resource (i.e., transmission time) is shared equally among flows unless this allocation fails to satisfy some class' rate requirement. In the latter case the throttled classes are allocated their minimum required rate ($x_i\varphi_i$) and the remaining capacity is equally shared among the flows of non-throttled classes. Hence for $\boldsymbol{x}\in\mathcal{S}$, $\tau_i(\boldsymbol{x})$ can be computed for classes in descending order as follows

$$\tau_K(\boldsymbol{x})=\frac{x_K}{\min\left(\sum_{i=1}^{K}x_i,\beta_K\right)},$$

$$\tau_i(\boldsymbol{x})=\max\left(\frac{x_i}{\beta_i},\frac{x_i}{\sum_{j=1}^{i}x_j}\left(1-\sum_{j=i+1}^{K}\tau_j(\boldsymbol{x})\right)\right).$$

2. Let i be a class such that all classes with lower feasible rates have no active flows, then, if accepting one more flow of this class leads to a feasible state, new flows are accepted with probability 1. In a more formal manner, for $i\ge\kappa(\boldsymbol{x})$ if $\boldsymbol{x}+\boldsymbol{e}_i\in\mathcal{S}$ then $p_i(\boldsymbol{x})=1$. Obviously, whatever the traffic class i, if $\boldsymbol{x}+\boldsymbol{e}_i\notin\mathcal{S}$, $p_i(\boldsymbol{x})=0$.

The first principle precisely specifies the BWA whereas the second one gives the AC probabilities only in some cases. Those cases not covered can be worked out by applying the fact that, since the system satisfies the detailed balance equations, it is reversible and, in particular, satisfies the *Kolmogorov's criterion* (see, for instance, [24, Chapter 10]). The method for doing so is detailed in what follows.

Through all discussion we assume that $\boldsymbol{x},\boldsymbol{x}+\boldsymbol{e}_i\in\mathcal{S}$ and $i<\kappa(\boldsymbol{x})$, otherwise the value of $p_i(\boldsymbol{x})$ is already known: $p_i(\boldsymbol{x})=0$ if $\boldsymbol{x}+\boldsymbol{e}_i\notin\mathcal{S}$, and $p_i(\boldsymbol{x})=1$ if $\boldsymbol{x}+\boldsymbol{e}_i\in\mathcal{S}$ but $i\ge\kappa(\boldsymbol{x})$.

Define $\xi_i(\boldsymbol{x})=\psi_i(\boldsymbol{x}+\boldsymbol{e}_i)/\rho_i=p_i(\boldsymbol{x})/\tau_i(\boldsymbol{x}+\boldsymbol{e}_i)$ and by applying the Kolmogorov's criterion to the cycle

$$\begin{tikzcd}
\boldsymbol{x}-x_{\kappa(\boldsymbol{x})}\boldsymbol{e}_{\kappa(\boldsymbol{x})}+\boldsymbol{e}_i \arrow[d] & \cdots \arrow[l] & \boldsymbol{x}+\boldsymbol{e}_i \arrow[l] \\
\boldsymbol{x}-x_{\kappa(\boldsymbol{x})}\boldsymbol{e}_{\kappa(\boldsymbol{x})} \arrow[r] & \cdots \arrow[r] & \boldsymbol{x} \arrow[u]
\end{tikzcd},$$

we obtain

$$\xi_i(\boldsymbol{x})=\xi_i(\boldsymbol{x}-x_{\kappa(\boldsymbol{x})}\boldsymbol{e}_{\kappa(\boldsymbol{x})})\prod_{j=1}^{x_{\kappa(\boldsymbol{x})}}\frac{\xi_{\kappa(\boldsymbol{x})}\left(\boldsymbol{x}-j\boldsymbol{e}_{\kappa(\boldsymbol{x})}+\boldsymbol{e}_i\right)}{\xi_{\kappa(\boldsymbol{x})}\left(\boldsymbol{x}-j\boldsymbol{e}_{\kappa(\boldsymbol{x})}\right)}. \tag{12}$$

Since

$$\xi_{\kappa(\boldsymbol{x})}(\boldsymbol{x})=\frac{1}{\tau_{\kappa(\boldsymbol{x})}(\boldsymbol{x}+\boldsymbol{e}_{\kappa(\boldsymbol{x})})}=\frac{1}{\max\left(\frac{x_{\kappa(\boldsymbol{x})}+1}{\beta_{\kappa(\boldsymbol{x})}},\frac{x_{\kappa(\boldsymbol{x})}+1}{1+\sum_{m=1}^{K}x_m}\right)}$$
$$=\frac{\min\left(\beta_{\kappa(\boldsymbol{x})},1+\sum_{m=1}^{\kappa(\boldsymbol{x})}x_m\right)}{x_{\kappa(\boldsymbol{x})}+1},$$

Eq. (12) becomes

$$\xi_i(\boldsymbol{x})=\xi_i(\boldsymbol{x}-x_{\kappa(\boldsymbol{x})}\boldsymbol{e}_{\kappa(\boldsymbol{x})})\frac{\min\left(\beta_{\kappa(\boldsymbol{x})},1+\sum_{m=1}^{\kappa(\boldsymbol{x})}x_m\right)}{\min\left(\beta_{\kappa(\boldsymbol{x})},1+\sum_{m=1}^{\kappa(\boldsymbol{x})-1}x_m\right)}, \tag{13}$$

and by applying (13) recursively it follows that

$$\xi_i(\boldsymbol{x})=\frac{\min\left(\beta_i,1+\sum_{m=1}^{i}x_m\right)}{x_i+1}\times$$
$$\prod_{j=i+1}^{\kappa(\boldsymbol{x})}\frac{\min\left(\beta_j,1+\sum_{m=1}^{j}x_m\right)}{\min\left(\beta_j,1+\sum_{m=1}^{j-1}x_m\right)}.$$

Finally, $p_i(\boldsymbol{x})$ can be computed as $p_i(\boldsymbol{x})=\tau_i(\boldsymbol{x}+\boldsymbol{e}_i)\xi_i(\boldsymbol{x})$.

3.2 Policy Evolution

The first policy ω_1 can be considered, in a way, biased towards less-favored traffic classes in terms of feasible rate, which makes sense given the low load situation. As load increases, however, situation changes and optimal policy orientation shifts towards efficiency, limiting the access to the system of the more resource-consuming traffic classes. More precisely, we say that class i consumes more resources than class j if $\mu_iC_i<\mu_jC_j$. In other words, resource consumption of a class is measured as the flow mean sojourn time in the system considering there are no other active flows. Throughout this subsection it is assumed without loss of generality that $\mu_1C_1\ge\mu_2C_2\ge\dots\ge\mu_KC_K$, i.e., traffic classes are sorted in ascending order according to resource consumption. It has been found that starting with ω_1, a series of transformations T_i, which penalize the most resource consuming classes and favor the least resource consuming ones, are successively applied as load increases

$$\omega_1\xrightarrow{T_1}\omega_2\xrightarrow{T_2}\cdots\xrightarrow{T_{m-1}}\omega_m.$$

The last policy ω_m is at the opposite side of ω_1, i.e., all resources are reserved for class 1, which is the least resource consuming class:

$$p_i(\boldsymbol{x}) = \begin{cases} 1 & \text{if } i = 1 \text{ and } \boldsymbol{x} + \boldsymbol{e}_1 \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases},$$

$$\tau_i(\boldsymbol{x}) = \begin{cases} 1 & \text{if } i = 1 \text{ and } \boldsymbol{x} \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \qquad i = 1, \dots, K.$$

Next we describe the type of policy transformations T_i. Before doing so we need to introduce some additional notation. Let us define

$$\mathcal{R} := \{\boldsymbol{x} = (0, x_2, \dots, x_K) : \boldsymbol{x} \in \mathcal{S}\},$$

and introduce the order relation $\succ$ defined as follows: we say that $\boldsymbol{x} \succ \boldsymbol{y}$ if $x_j > y_j$ and $x_i = y_i$ for $i = j+1, \dots, K$. Now consider that $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{|\mathcal{R}|}$ is a sorted list of all the elements in $\mathcal{R}$, i.e. $\boldsymbol{y}_1 \succ \boldsymbol{y}_2 \succ \cdots \succ \boldsymbol{y}_{|\mathcal{R}|}$. Finally, for each $\boldsymbol{y}_i$ we define the set

$$\Delta_i := \{\boldsymbol{x} = x_1\boldsymbol{e}_1 + \boldsymbol{y}_i, \boldsymbol{x} \in \mathcal{S}\}.$$

Let us start with the policy ω_1. The set of feasible states for ω_1 is $\mathcal{S}_1 = \mathcal{S}$. The transformation T_1 will affect one or more states in Δ_1 in one of the following ways:

A: the bandwidth allocation to class K is set to its minimum, i.e., $\tau_K(\boldsymbol{x}) = \beta_K x_K$, and the released capacity is shared by the remaining classes

B: some admission probabilities of the least resource consuming class are set to 1 (if all the admission probabilities in Δ_1 of the least resource consuming class are already 1, the second least resource consuming class is considered and so on)

C: states in Δ_1 are made unfeasible by rejecting those flow arrival that would lead the system to a state in Δ_1, i.e., if $\boldsymbol{x} + \boldsymbol{e}_K \in \Delta_1$ then $p_K(\boldsymbol{x}) = 0$.

Note that since the detailed balanced condition has to be satisfied, changes applied to a state may also affect other neighboring states, which might be outside Δ_1. Before the type-C transformation is applied, none or several transformations of types A or B can be applied. Obviously after the type-C transformation no more transformations can target states in Δ_1 since these are not feasible anymore. After the type-C transformation, the set of feasible states becomes $\mathcal{S}_2 = \mathcal{S}_1 \setminus \Delta_1$, then none or several type-A,B transformations are applied to states in Δ_2 followed by the type-C transformation which clips the feasible state space to $\mathcal{S}_3 = \mathcal{S}_2 \setminus \Delta_2$,... This process is repeated until the feasible state space becomes $\mathcal{S}_M := \{(x_1, 0, \dots, 0) : x_1 \leq \beta_1\}$, which corresponds to the last policy ω_m. Note that after some type-C transformations (more specifically after $\lfloor \beta_K \rfloor$ of them) no class-K flows are let into the system and class $(K-1)$ will then play the role of the most resource consuming class. Again, when class $(K-1)$ has been completely removed it will be substituted by class $(K-2)$, and so forth until only class 1 is let into the system.

3.3 Suboptimal Policies

The description of policy transformations given in previous section is not sufficient to obtain $\omega_{j\geq 2}$ from ω_1. That will require, at least, knowing how many and in which order type-A,B transformations occur between two type-C transformations. Unfortunately, in our experiments we could not observe any general and simple rule for the occurrence of transformations of types A and B. Besides, even if we were able to determine the exact sequence of transformations the load values at which transformations occur ($\rho^{(j)}$) will remain unknown. As mentioned above, not knowing the thresholds $\rho^{(j)}$ can be circumvented by the approach of Eq. (11) but this requires computing P_b for each policy ω_j. Observe that if the system state probabilities have been computed under policy ω_j and T_j is of type C, the system state probabilities under policy ω_{j+1} can be recomputed by simply renormalizing.

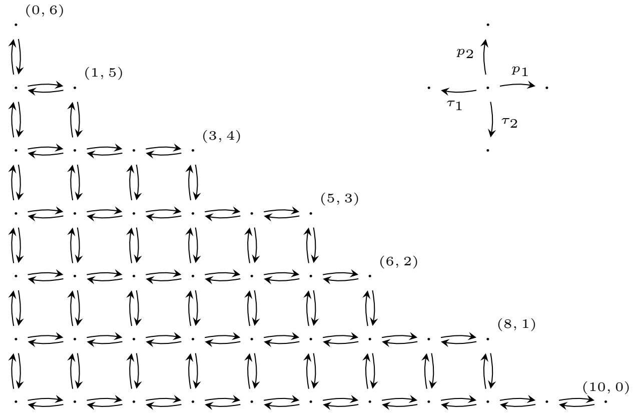


Figure 1: State space.

Motivated by the aforementioned reasons we propose a set of suboptimal policies which are defined as follows: $\hat{\omega}_1 \equiv \omega_1$ and $\hat{\omega}_j \equiv \omega_1(\text{restricted to } \mathcal{S}_j)$ for $j = 2, \dots, M$, which is equivalent to say $\hat{\omega}_1 \equiv \omega_1$ and then only the type-C transformations are applied. By definition $\hat{\omega}_1 = \omega_1$, and also $\hat{\omega}_M = \omega_m$ but in general $\hat{\omega}_j$ is not necessarily included in $\{\omega_1, \dots, \omega_m\}$ since, as noted previously, a transformation of type B or C may also affect states outside its target set of states Δ_l. For a given value of ρ the suboptimal policy can be obtained using the approach of Eq. (11), $\hat{\omega}(\rho) = \arg\min_{\hat{\omega}_j} P_b(\hat{\omega}_j, \rho)$.

In the next section we present a numerical evaluation example that shows the good performance achieved by the suboptimal policies introduced here.

4. NUMERICAL EXAMPLES

Consider the basic configuration: $(C_1, C_2) = (5, 3)$; $(\varphi_1, \varphi_2) = (1/2, 1/2)$; $(\mu_1, \mu_2) = (1, 1)$; $\hat{\boldsymbol{\rho}} = (2/5, 3/5)$; Fig. 1 displays its state space.

4.1 First Policy

The first policy ω_1, which is obtained as described in Section 3.1, is given by

$$[\tau_1(i,j)]_{ij} = \begin{bmatrix} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1/2 & 1/3 & 1/4 & 1/5 & 1/6 & \cdot \\ 1 & 2/3 & 1/2 & 2/5 & 1/3 & \cdot & \cdot \\ 1 & 3/4 & 3/5 & 1/2 & 1/3 & \cdot & \cdot \\ 1 & 4/5 & 2/3 & 1/2 & \cdot & \cdot & \cdot \\ 1 & 5/6 & 2/3 & 1/2 & \cdot & \cdot & \cdot \\ 1 & 5/6 & 2/3 & \cdot & \cdot & \cdot & \cdot \\ 1 & 5/6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 5/6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix},$$

$$[\tau_2(i,j)]_{ij} = \begin{bmatrix} & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1/2 & 2/3 & 3/4 & 4/5 & 5/6 & \cdot \\ 0 & 1/3 & 1/2 & 3/5 & 2/3 & \cdot & \cdot \\ 0 & 1/4 & 2/5 & 1/2 & 2/3 & \cdot & \cdot \\ 0 & 1/5 & 1/3 & 1/2 & \cdot & \cdot & \cdot \\ 0 & 1/6 & 1/3 & 1/2 & \cdot & \cdot & \cdot \\ 0 & 1/6 & 1/3 & \cdot & \cdot & \cdot & \cdot \\ 0 & 1/6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1/6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix},$$

$$[p_1(i,j)]_{ij} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \cdot \\ 1 & 1 & 1 & 1 & 2/3 & \cdot & \cdot \\ 1 & 1 & 1 & 3/4 & 0 & \cdot & \cdot \\ 1 & 1 & 4/5 & 3/5 & \cdot & \cdot & \cdot \\ 1 & 5/6 & 2/3 & 0 & \cdot & \cdot & \cdot \\ 1 & 5/6 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & 5/6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix},$$

$$[p_2(i,j)]_{ij} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & \cdot \\ 1 & 1 & 1 & 1 & 1 & \cdot & \cdot \\ 1 & 1 & 1 & 1 & 0 & \cdot & \cdot \\ 1 & 1 & 1 & 1 & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & 0 & \cdot & \cdot & \cdot \\ 1 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}.$$

4.2 Policy Evolution

Figures 2–7 show the evolution of the policy parameters as the load increases. Figure 8(a) depicts the blocking probability and a "summary" of the policy evolution. For each value of x_2 the admission probabilities for class 2, $p_2(x_1, x_2)$, have been averaged over those values of x_1 such that $(x_1, x_2{+}1) \in \mathcal{S}$, i.e., $p_2(x_1, x_2) > 0$ in ω_1. The resulting curves show the relative position (loadwise) of policy changes affecting a "row" of states (x_2 constant), and in particular values of ρ at which such rows are removed from the feasible states.

Figure 8(b) shows the same type of plot as Fig. 8(a) but now $\hat{\boldsymbol{\rho}} = (1/3, 2/3)$ has been varied. From the shape of the curves we observe that, as expected, varying $\hat{\boldsymbol{\rho}}$ changes the values $\rho^{(j)}$ but not the set of optimal policies ω_j.

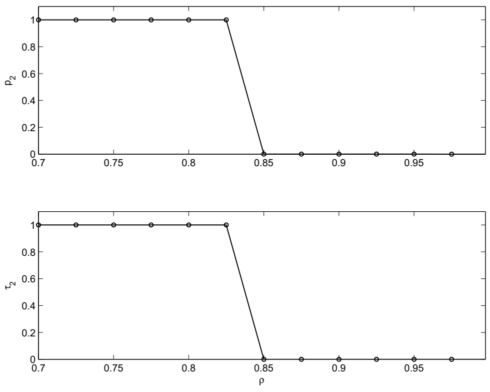


Figure 2: $p_2(0,5), \tau_2(0,6)$

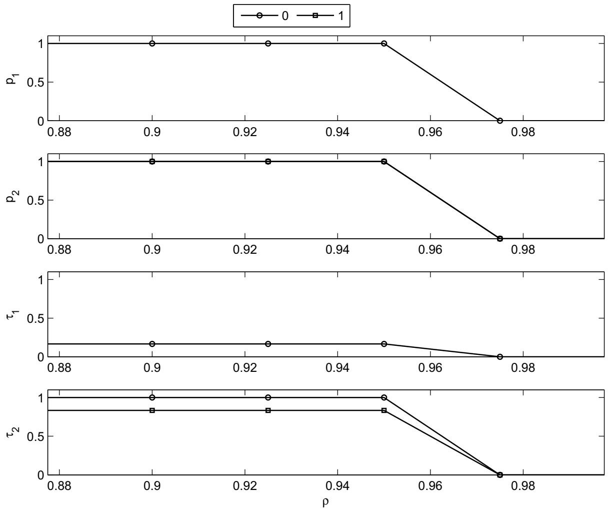


Figure 3: $p_1(x_1,5), p_2(x_1,4), \tau_1(x_1{+}1,5), \tau_2(x_1,5)$**; curves are parameterized by** x_1

In order to see the effect of modifying the resource consumption ordering we set $\mu_2 = 2$. Now $C_1 = 5 > 3 = C_2$ but $\mu_1 C_1 = 5 < 6 = \mu_2 C_2$, so the optimal policy evolves limiting the access of class 1 traffic as shown in Fig. 9.

4.3 Comparison of policies

The curves in Fig. 10 represent the relative value of P_b for the different policies taking the optimal insensitive policy as the reference. The first policy (ω_1) show important degradations as the load moves away from their optimality regions so it does not seem advisable to keep using ω_1 far beyond $\rho^{(1)}$. It is noticeable that the suboptimal policy $\hat{\omega}$ is an excellent approximation to ω, which is the targeted optimum, so its relative performance is also excellent; in this scenario the maximum deviation of $P_b(\hat{\omega})$ from $P_b(\omega)$ is a 1.3%. We also plotted a curve corresponding to the optimal (non-necessarily insensitive) policy, which exhibits an important gain over the more restrictive case of insensitive policies. For this curve, though, the validity of the results is limited to the case where the flow sizes are exponentially distributed. In order to compute P_b for the optimal policy the set of equations corresponding to the detailed balanced con-

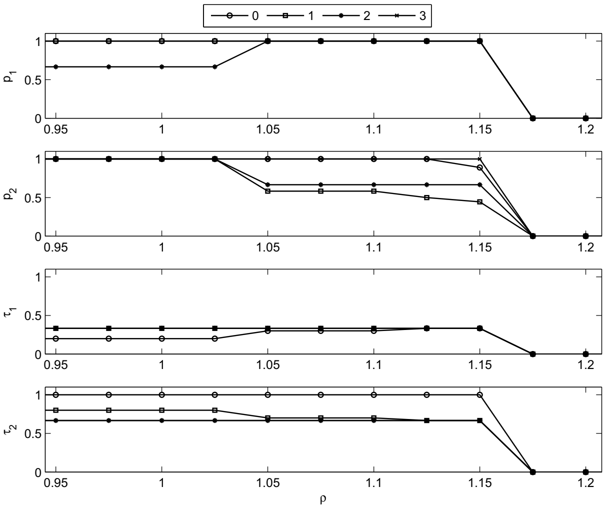


Figure 4: $p_1(x_1, 4), p_2(x_1, 3), \tau_1(x_1+1, 4), \tau_2(x_1, 4)$**; curves are parameterized by** x_1

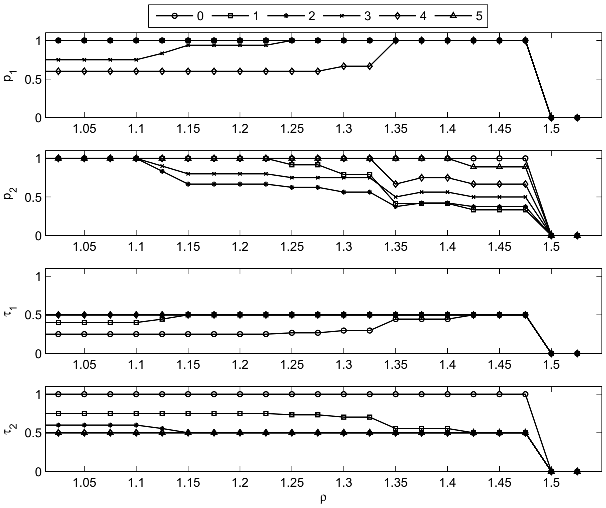


Figure 5: $p_1(x_1, 3), p_2(x_1, 2), \tau_1(x_1+1, 3), \tau_2(x_1, 3)$**; curves are parameterized by** x_1

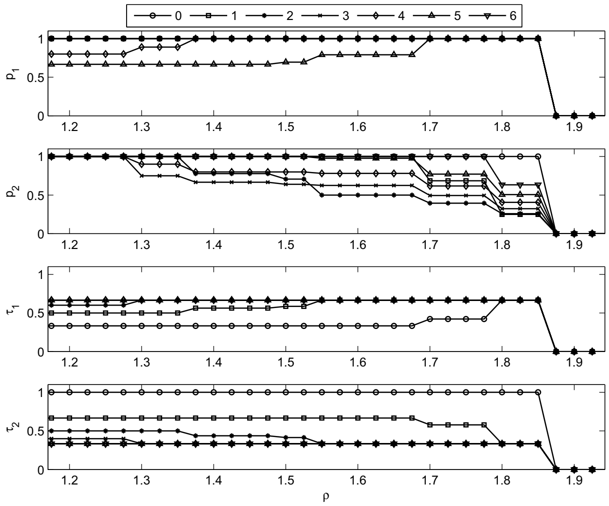


Figure 6: $p_1(x_1, 2), p_2(x_1, 1), \tau_1(x_1+1, 2), \tau_2(x_1, 2)$**; curves are parameterized by** x_1

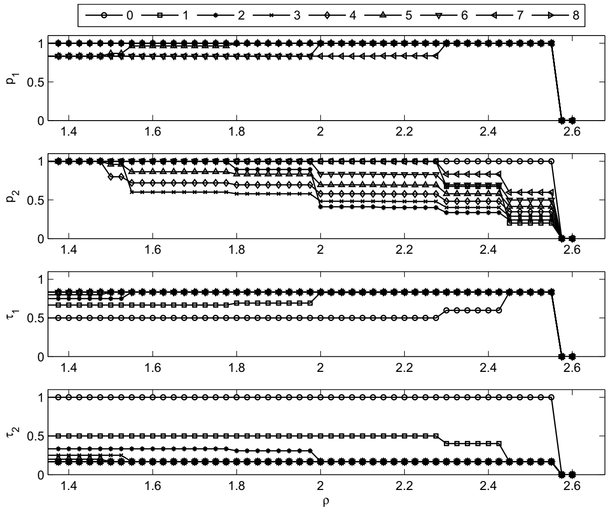
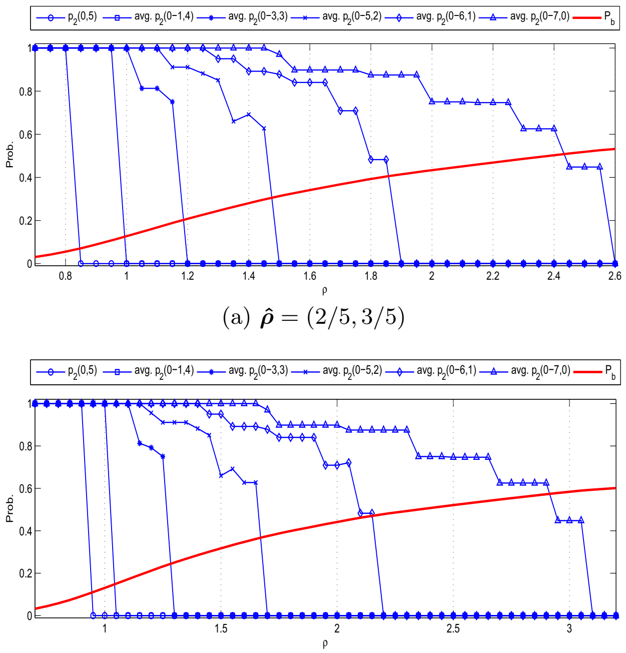


Figure 7: $p_1(x_1, 1), p_2(x_1, 0), \tau_1(x_1+1, 1), \tau_2(x_1, 1)$**; curves are parameterized by** x_1

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376



(a) $\hat{\boldsymbol{\rho}} = (2/5, 3/5)$

(b) $\hat{\boldsymbol{\rho}} = (1/3, 2/3)$

Figure 8: Blocking probability and policy variations

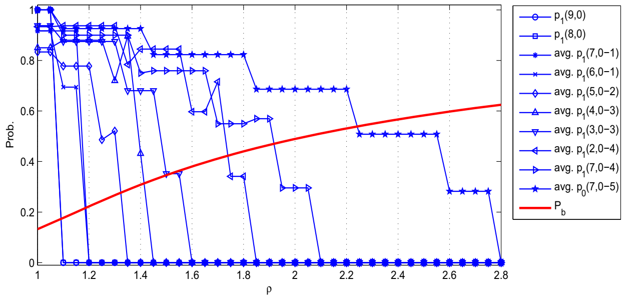


Figure 9: Blocking probability and policy variations; $\mu_2 = 2$

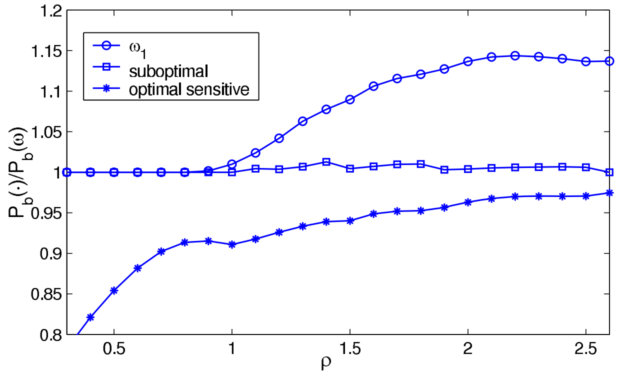


Figure 10: Relative performance: basic configuration.

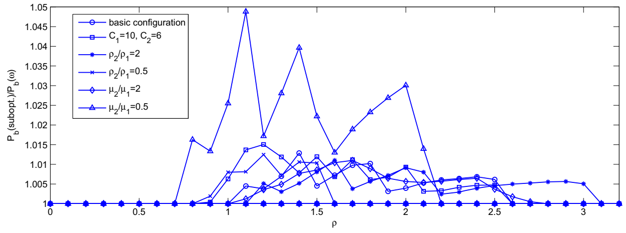


Figure 11: Relative performance of $\hat{\omega}$: sensitivity to configuration parameters.

dition in the linear program (see Eq. (8)) were substituted by the global balance equations

$$\sum_i \Bigg(I_{\{\boldsymbol{x}+\boldsymbol{e}_i \in \mathcal{S}\}} \lambda_i \sum_{\boldsymbol{d}:d_i=1} \sum_b \pi(\boldsymbol{x}, \boldsymbol{d}, b) \\ + I_{\{\boldsymbol{x}-\boldsymbol{e}_i \in \mathcal{S}\}} C_i \mu_i \sum_{\boldsymbol{d}} \pi(\boldsymbol{x}, \boldsymbol{d}, i) \\ - \lambda_i \sum_{\boldsymbol{d}:d_i=1} \sum_b \pi(\boldsymbol{x}-\boldsymbol{e}_i, \boldsymbol{d}, b) \\ - C_i \mu_i \sum_{\boldsymbol{d}} \pi(\boldsymbol{x}+\boldsymbol{e}_i, \boldsymbol{d}, i) \Bigg) = 0 \quad \forall \boldsymbol{x} \in \mathcal{S},$$

where $I_{\{\cdot\}}$ is the indicator function and by convention $\pi(\boldsymbol{x}, \boldsymbol{d}, b) = 0$ if $\boldsymbol{x} \notin \mathcal{S}$.

In Fig. 11 the sensitivity of the relative suboptimal performance to different configuration parameters is analyzed. All curves but the last one ($\mu_2/\mu_1 = 0.5$) display an excellent performance of $\hat{\omega}$. Further experiments in that direction revealed that it is indeed the imbalance between $\mu_1 C_1$ and $\mu_2 C_2$ that is the cause of the performance degradation.

5. CONCLUSION

We have considered the joint optimization of bandwidth allocation and admission control for elastic users competing for a downlink data channel in a cellular network. Robustness and generality of the results were main concerns in our research and so we focused on those strategies that are insensitive to the detailed traffic characteristics beyond mean values. The optimization problem has been formulated using a *Markov Decision Process-Linear Programming* approach. A characterization of the optimal policy has been obtained inductively. It has been found that the optimal policy is a piecewise constant function of the system load having only finitely many pieces. Moreover, the policy settings for each piece do only depend on the minimum rate requirements and feasible rates, in particular they are not dependent on the arrival rates. These features confer additional robustness to the solution.

We observed that except for low loads the complexity of computing the optimal policy may make this policy impractical. As an alternative we proposed a much simpler suboptimal policy that satisfies the same requirements and achieves a good relative performance unless the values $\mu_i C_i$ (reciprocal of the mean sojourn time if a class-i user was alone in the system) for the different user classes are significantly imbalanced.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376

Acknowledgments

A major part of this work was done while Vicent Pla was visiting the Networking Laboratory at Helsinki University of Technology. This author gratefully acknowledges the financial support of EuroNGI NoE. He also expresses his gratitude to Pasi Lassila and Aleksi Penttinen for their help, fruitful discussions and excellent work climate, and to Sarah Bazzocco and Jesús Planagumà for their friendship.

6. REFERENCES

[1] S. Aalto, U. Ayesta, S. Borst, V. Misra, and R. Núñez-Queija. Beyond processor sharing. *SIGMETRICS Performance Evaluation Review*, 34(4):36–43, 2007.

[2] E. Altman, K. Avrachenkov, and U. Ayesta. A survey on discriminatory processor sharing. *Queueing Systems*, 53(1):53–63, 2006.

[3] K. Avrachenkov, U. Ayesta, P. Brown, and R. Núñez-Queija. Discriminatory processor sharing revisited. In *Proceedings of IEEE INFOCOM 2005*, volume 2, pages 784–795, 2005.

[4] T. Bonald. Insensitive queueing models for communication networks. In *Valuetools '06: Proceedings of the 1st international conference on Performance evaluation methodolgies and tools*, page 57, New York, NY, USA, 2006. ACM Press.

[5] T. Bonald, S. Borst, N. Hegde, and A. Proutière. Wireless data performance in multi-cell scenarios. In *SIGMETRICS '04/Performance '04: Proceedings of the joint international conference on Measurement and modeling of computer systems*, pages 378–380, New York, NY, USA, 2004. ACM.

[6] T. Bonald, S. C. Borst, and A. Proutière. How mobility impacts the flow-level performance of wireless data systems. In *INFOCOM 2004*, volume 3, pages 1872–1881. IEEE, 2004.

[7] T. Bonald, M. Jonckheere, and A. Proutière. Insensitive load balancing. In *SIGMETRICS '04/Performance '04: Proceedings of the joint international conference on Measurement and modeling of computer systems*, pages 367–377, New York, NY, USA, 2004. ACM Press.

[8] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing Systems: Theory and Applications*, 53(1-2):65–84, 2006.

[9] T. Bonald and A. Proutière. Insensitivity in processor-sharing networks. *Performance Evaluation*, 49(1-4):193–209, 2002.

[10] T. Bonald and A. Proutière. Insensitive bandwidth sharing in data networks. *Queueing Systems: Theory and Applications*, 44(1):69–100, 2003.

[11] T. Bonald and A. Proutière. Wireless downlink data channels: user performance and cell dimensioning. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 339–352, New York, NY, USA, 2003. ACM Press.

[12] T. Bonald and J. W. Roberts. Congestion at flow level and the impact of user behaviour. *Computer Networks*, 42:521–536, 2003.

[13] T. Bonald and J. W. Roberts. Scheduling network traffic. *SIGMETRICS Performance Evaluation Review*, 34(4):29–35, 2007.

[14] S. Borst. User-level performance of channel-aware scheduling algorithms in wireless data networks. *IEEE/ACM Transactions on Networking*, 13(3):33–47, June 2005.

[15] S. Borst and N. Hegde. Integration of streaming and elastic traffic in wireless networks. In *INFOCOM 2007*, pages 1884–1892. IEEE, 2007.

[16] S. Borst, A. Proutière, and N. Hegde. Capacity of Wireless Data Networks with Intra-and Inter-Cell Mobility. In *INFOCOM 2006*, pages 1–12. IEEE, 2006.

[17] S. C. Borst, K. L. Clarkson, J. M. Graybeal, H. Viswanathan, and P. A. Whiting. User-level QoS and traffic engineering for 3G wireless 1xEV-DO systems. *Bell Labs Technical Journal*, 8(2):33–47, 2003.

[18] G. Fayolle, I. Mitrani, and R. Iasnogorodski. Sharing a processor among many job classes. *J. ACM*, 27(3):519–532, 1980.

[19] M. Jonckheere and J. Virtamo. Optimal insensitive routing and bandwidth sharing in simple data networks. In *SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 193–204, New York, NY, USA, 2005. ACM Press.

[20] F. P. Kelly. *Reversibility and Stochastic Networks.* Wiley New York, 1979.

[21] J. Leino and J. Virtamo. Insensitive load balancing in data networks. *Computer Networks*, 50(8):1059–1068, June 2006.

[22] S. Liu and J. Virtamo. Performance analysis of wireless data systems with a finite population of mobile users. In *Proceedings of the 19th International Teletraffic Congress ITC 19*, pages 1295–1304, 2005.

[23] L. Massoulié and J. W. Roberts. Arguments in favour of admission control for TCP flows. In *Proceedings of ITC 16*, 1999.

[24] R. Nelson. *Probability, Stochastic Processes and Queueing Theory.* Springer-Verlag, 1995.

[25] R. Núñez-Queija and H.-P. Tan. Location-based admission control for differentiated services in 3G cellular networks. In *Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems (MSWiM'06)*, pages 322–329, New York, NY, USA, 2006. ACM.

[26] J. W. Roberts. Traffic theory and the Internet. *Communications Magazine, IEEE*, 39(1):94–99, 2001.

[27] S. M. Ross. *Applied probability models with optimization applications.* Holden-Day, 1970.

[28] Y. Wu, C. Williamson, and J. Luo. On processor sharing and its applications to cellular data network provisioning. *Performance Evaluation*, 64(9–12):892–908, Oct. 2007.

[29] S. F. Yashkov and A. S. Yashkova. Processor sharing: A survey of the mathematical theory. *Autom. Remote Control*, 68(9):1662–1731, 2007.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4376
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4376