Application-Oriented Flow Control: Fundamentals, Algorithms and Fairness

Wei-Hua Wang, Marimuthu Palaniswami, *Senior Member, IEEE*, and Steven H. Low, *Senior Member, IEEE*

Abstract—This paper is concerned with flow control and resource allocation problems in computer networks in which real-time applications may have hard quality of service (QoS) requirements. Recent optimal flow control approaches are unable to deal with these problems since QoS utility functions generally do not satisfy the strict concavity condition in real-time applications. For elastic traffic, we show that bandwidth allocations using the existing optimal flow control strategy can be quite unfair. If we consider different QoS requirements among network users, it may be undesirable to allocate bandwidth simply according to the traditional max-min fairness or proportional fairness. Instead, a network should have the ability to allocate bandwidth resources to various users, addressing their real utility requirements. For these reasons, this paper proposes a new distributed flow control algorithm for multiservice networks, where the application's utility is only assumed to be continuously increasing over the available bandwidth. In this, we show that the algorithm converges, and that at convergence, the utility achieved by each application is well balanced in a proportionally (or max-min) fair manner.

Index Terms—Congestion control, quality of service, real-time application, resource allocation, utility max-min fairness, utility proportional fairness.

I. INTRODUCTION

WITH THE RAPID growth and development of computer technologies, traditional telephone and television networks are gradually being merged into computer networks. It is now desirable to transmit high quality real-time multimedia information through one multiservice network, such as in the world widely deployed Internet.

Using *implicit* congestion information, packet loss and transmission delay, Internet Transmission Control Protocol (TCP) is able to provide most efficient services for the current *elastic* [1] data applications like web browsing, file transference and electronic mail. But it is not sufficient to support real-time applications, such as teleconferencing, audio and video broadcasting, etc. These applications are generally delay sensitive and have strict quality of service (QoS) requirements. Unlike conventional data transmissions, in real-time applications, the lack of bandwidth may severely affect QoS performance.

To provide efficient congestion control for the Internet and in other communication networks, like *explicit* rate-based ATM available bit rate (ABR) services, a new systematic method called "Optimal Flow Control" (OFC) has been recently proposed by Kelly [2]. In [2], the network flow control problem is formulated, for the first time, as an optimization problem and explicit rate flow control algorithm is derived by solving the optimization problem via a link congestion pricing policy. Following this pioneering work, an extensive study of network flow control based on optimization method and game theory has been carried out by network researchers, e.g., [3]–[13] for single-path networks, [14], [15] for multiple-path networks and [16], [17] for multirate multicast networks.

Even though many authors use different mathematical formulations and a variety of optimization techniques, their approaches are essentially the same in the literature. In matters of optimal flow control, the main idea is that for every network user (source), there is an associated utility function which may be used to a measurement of the application's QoS performance over the available bandwidth. The objective of OFC is to maximize the aggregate utility (QoS performance) for all users under the link capacity constraints across the network. An optimal flow control algorithm is then derived by solving the optimization problem distributively. It usually consists of a link algorithm that measures the congestion link price in the network and a source algorithm that adapts the transmission rate in response to congestion feedback signals. At convergence, traffic in the network is brought to an optimal operating point. Meanwhile, congestion, like buffer backlog at each link, is well maintained and controlled.

Optimal flow control not only provides an efficient congestion control mechanism for the network, it also gives an easy way to provide a fair bandwidth allocation among competing users. By selecting a logarithmic function for utility, Kelly [2] showed that the optimal flow control approach achieves (in equilibrium) a *proportional fairness* in bandwidth allocation. Using a standard optimal flow control formulation, an additional fairness criterion called *max-min fairness* [18] has been investigated by Mo and Walrand [6] and Marbach [19]. Mo and Walrand [6] used a family of utility functions to approximate an arbitrarily close max-min fair allocation, and Marbach [19] proved the existence of a (weight) max-min fair allocation in an optimal flow controlled priority service network.

Optimal flow control approach has also been used to analyze current major TCP congestion control protocols, such as

W.-H. Wang and M. Palaniswami are with the Department of Electrical and Electronic Engineering, University of Melbourne, Victoria 3010, Australia (e-mail: weihw@ee.unimelb.edu.au; swami@ee.unimelb.edu.au).

S. H. Low is with the Departments of Computer Science and Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: slow@caltech.edu).

Reno [20], Vegas [21], RED [22], and REM [23]. Within the OFC framework, we are able to see that various TCP protocols are merely different algorithms to solve the same optimization problem with different source utility functions [8], [24].

Despite great advances in optimal flow control theory and its application, serious limitations still exist.

- At present, optimal flow control approach is only suitable for *elastic* traffic, which has an increasing and strictly concave utility function that ensures the feasibility of optimal solutions and the convergence of flow control algorithms. Optimal flow control is not capable of dealing with congestion, or resource allocations in computer networks, when real-time applications are engaged.
- Sometimes negative utility values do not give a clear understanding for practical engineering, although they are commonly used in logarithmic utility functions.
- In practical Internet engineering, the utility of an application may not be given by an explicit function, except that a numerical description may be available. Thus, the derivative of such a utility function, which is used by flow control algorithms, cannot be measured accurately.
- There exists a serious conflict between the QoS balance and the utility maximization. If users select utility functions based on their real QoS requirements, then the optimal flow control approach will result in a totally unfair resource allocation within the network. In particular, an application with low bandwidth demand is likely to receive a high bandwidth allocation, and vice versa.

Many of the above issues will be addressed in the following section.

In this paper, we revisit the flow control problem as it exists in computer networks, and give special consideration to real-time applications which are sensitive to transmission bandwidth. Here the application's utility function is assumed to be strictly increasing, but may not be strictly concave. We propose a new distributed flow control algorithm and its convergence is proven. Moreover, at convergence, we show that the bandwidth is allocated properly within the network and the utility achieved by each source is in the proportional fairness.

The rest of the paper is organized as follows. In Section II, we review the background of flow control problem and examine the recently proposed optimal flow control approach. In Section III, we formulate the application-oriented flow control problem and propose a new resource allocation criterion which we call "utility proportional fairness." In Section IV, we develop a new flow control algorithm in order to achieve utility proportional fairness, as well as an alternative algorithm in order to obtain utility max-min fairness. Finally, we present numerical results to illustrate the performance of our algorithms in Section V and draw conclusions in Section VI.

II. BACKGROUND

In this section, we give a brief introduction to the problem of dynamic resource allocation and congestion control in computer networks. Then we review the recently proposed optimal flow control approach and address in detail its limitations.

A. Network Flow Control Problem

Consider a network that consists of a set $\mathcal{L} = \{1, 2, \ldots, L\}$ of unidirectional links with a capacity c_l, $l \in \mathcal{L}$. The network is shared by a set $\mathcal{S} = \{1, 2, \ldots, S\}$ of sources. Each source s is characterized by five parameters $\{\mathcal{L}_s, x_s, m_s, M_s, U_s(x_s)\}$. The path $\mathcal{L}_s \subseteq \mathcal{L}$ is a subset of links that connect each source s to its destination, x_s is the transmission rate satisfying $0 \leq m_s \leq x_s \leq M_s < \infty$, where m_s and M_s are the minimum and maximum transmission rates required by source s, respectively. $U_s(x_s) : \mathcal{R}^+ \mapsto \mathcal{R}$ is a continuously increasing and bounded utility function which can be used as a QoS performance indicator for source s.

Let $x = [x_1, x_2, \ldots, x_S]^T$. For each link l, define $\mathcal{S}_l = \{s \in \mathcal{S} | l \in \mathcal{L}_s\}$, which is the set of sources that access link l. Note that, $l \in \mathcal{L}_s$ if and only if $s \in \mathcal{S}_l$.

In order to formulate the network flow control problem, we first define the notion of feasible (or attainable) bandwidth allocation.

Definition 1: A bandwidth allocation $x = [x_1, x_2, \ldots, x_S]^T$ is *feasible* or *attainable* if and only if $x_s \in [m_s, M_s]$, and no links in the network are congested, i.e.,

$$\sum_{s \in \mathcal{S}_l} x_s \leq c_l \quad \forall l \in \mathcal{L}. \qquad (1)$$

Throughout the paper, we assume that a minimum allocation $x = [m_1, m_2, \ldots, m_S]^T$ is attainable in the network.

The major task in network flow control is to guide traffic to a feasible bandwidth allocation, in such a way that each source is treated in a fair manner and guaranteed high utility performance.

When network bandwidth is abundant, there is no difficulty in satisfying every source utility, i.e., if $x = [M_1, M_2, \ldots, M_S]^T$ is attainable within the network. If the bandwidth is not sufficient (or even worse is scarce), then there arises a problem of how to allocate the existing resource fairly among competing sources that have different QoS requirements.

A simple example may be helpful in understanding this critical situation. Suppose there is a single link of bandwidth 3 shared by two sources. Source 1 attains a utility of $x_1/2$ with a maximum bandwidth requirement of 2, source 2 attains a utility of $x_2/4$ with a maximum bandwidth requirement of 4. A simple way to allocate bandwidth is by a factor of 3/2 to each source, with the aim of equally sharing and this is called *max-min fairness* [25]. If we consider the different bandwidth requirements these two sources might have, then it may be appropriate to allocate $x_1 = 1$ and $x_2 = 2$, so that they both attain a utility of 1/2. However, this is not consistent with the standard optimal flow control approach, which when applied, will result in $x_1 = 2$ and $x_2 = 1$ with a maximum total utility of 5/4. In this situation, it is obvious that source 1 is much favored by a utility of 1 and source 2 is poorly treated with a utility of 1/4.

We will now show clearly why such unfairness can arise in the utility maximization-based optimal flow control approach.

B. Optimal Flow Control Approach

In [2], [3], and [5], the optimal flow control theorists try to find an "optimal" resource allocation so that the sum of all the source utilities is maximized, i.e., to pursue maximum of social

welfare in a network. This objective can be achieved by solving the following primal optimization problem:

$$\mathbf{P} : \max_{m_s \leq x_s \leq M_s} \quad U(x) = \sum_{s=1}^{S} U_s(x_s) \tag{2}$$

$$\text{subject to} \quad \sum_{s \in \mathcal{S}_l} x_s \leq c_l, \quad l = 1, \ldots, L. \tag{3}$$

Under a more critical assumption of strict concavity on utility functions, there always exists a unique optimal solution x to the maximization problem $\mathbf{P}$. The optimal solution x can be obtained by looking for a *saddle-point* in the following Lagrangian form:

$$L(x,p) = U(x) - \sum_{l=1}^{L} p_l \left(\sum_{s \in \mathcal{S}_l} x_s - c_l \right) \tag{4}$$

$$= \sum_{s=1}^{S} \left[U_s(x_s) - x_s \sum_{l \in \mathcal{L}_s} p_l \right] + \sum_{l=1}^{L} p_l c_l \tag{5}$$

$$= \sum_{s=1}^{S} \left[U_s(x_s) - x_s p^s \right] + \sum_{l=1}^{L} p_l c_l \tag{6}$$

where the Lagrangian multipliers $p = [p_1, p_2, \ldots, p_L]^{\mathrm{T}} \geq 0$ are called link prices in the optimal flow control literature, and $p^s = \sum_{l \in \mathcal{L}_s} p_l$ is the path price for source s, which is the sum of the link prices along path $\mathcal{L}_s$.

An optimal flow control algorithm can be further extrapolated by solving the unconstrained max-min problem in the Lagrangian form $L(x,p)$:

$$\max_{m_s \leq x_s \leq M_s} \quad \min_{p \geq 0} \quad L(x,p). \tag{7}$$

Different optimization techniques used to solve the max-min problem above, may lead to different flow control algorithms. As an example, we present the following basic iterative algorithm using the dual method proposed by Low and Lapsley [5]:

$$x_s(t+1) = \arg \max_{m_s \leq x_s \leq M_s} L\left(x(t), p(t)\right)$$

$$= \left[U_s'^{-1} \left(p^s(t) \right) \right]_{m_s}^{M_s} \tag{8}$$

$$p_l(t+1) = \left[p_l(t) - \gamma \frac{\partial L}{\partial p_l} \left(x(t), p(t) \right) \right]^{+}$$

$$= \left[p_l(t) + \gamma \left(x^l(t) - c_l \right) \right]^{+} \tag{9}$$

where $\gamma > 0$ is a small step size, $[z]^{+} = \max\{0, z\}$, $[z]_a^b = \max\{a, \min\{b, z\}\}$, $x^l(t) = \sum_{s \in \mathcal{S}_l} x_s(t)$ is the aggregate source rate at link l. $U_s'^{-1}$ is the inverse of U_s' and is decreasing over the range $[U_s'(M_s), U_s'(m_s)]$, which attains maximum at the Lagrangian form $L(x,p)$, when p is fixed.

With the cooperation of source algorithm (8) and link algorithm (9), it has been shown in [5] that, when the step size γ is selected to be appropriately small, the algorithm converges at an optimal point that maximizes the total utilities within the network. This formulates a distributive framework in the optimal flow control strategy.

C. Conflict Between Utility Maximization and Bandwidth Fair Sharing

The optimal flow control approach is able to achieve the maximum aggregate utility in the system, however, a seriously unfair situation may still result in bandwidth allocation among competing users.

Consider an individual source s associated with utility $U_s(x_s) = a_s \log x_s$, where $a_s > 0$ is a weighting parameter. It is well known that in this case, an a_s-weighted proportional fairness [2], [3], [6] will be achieved at optimum (x^*, p^*), in which

$$U_s'(x_s^*) = \frac{a_s}{x_s^*} = p^{s^*} = \sum_{l \in \mathcal{L}_s} p_l^* \tag{10}$$

and

$$x_s^* p^{s^*} = x_s^* \sum_{l \in \mathcal{L}_s} p_l^* = a_s. \tag{11}$$

If we value each link resource as bandwidth $\times$ link_price, and not as bandwidth only,[1] (11) implies that in the weighted proportional fairness, each source is allocated a_s units of network resources, and that the total amount of network resources is $\sum_{s \in \mathcal{S}} a_s$.

In fact, a large a_s in the utility function means that the source is more easily satisfied with an increase in bandwidth and actually only requires a low bandwidth to achieve a high utility output. However, in the optimal flow control approach, the same source is always allocated a large amount of the network resource. Conversely, a source with a high bandwidth requirement (a small a_s in the utility function) is treated poorly, and is often allocated less bandwidth than required. (For a detailed illustration of bandwidth unfairness in OFC, please see Example 1 in Section V.)

This problem arises because the optimal flow control approach aims at maximizing the total utilities in the system. To achieve this objective, the optimal flow control policy always favors those sources in low demand (with a rapidly increasing utility), and is not willing to allocate bandwidth to sources in high demand, since such an allocation only contributes to a small increase in the aggregate utilities. This approach is efficient for congestion control and is fair in terms of bandwidth allocation only when all the sources attain the same (strictly concave and increasing) utility functions.

It is impractical to assume that all network applications have the same QoS requirements for bandwidth. Moreover, in real-time applications, their associated QoS utility may not satisfy the strict concavity condition. Furthermore, a variation in utility function by a constant amount may well make a significant difference in QoS performance in any practical application. However, in the utility maximization approach, there will be no change in the final bandwidth allocation. For example, the utility functions in the class $\log ax (a > 0)$ are treated exactly the same no matter how different a is.

For these reasons, our study is looking at QoS utilities in real-time applications, and we propose a new application-oriented

---

[1]This is more reasonable since different links may have different link prices due to different capacities and demands.
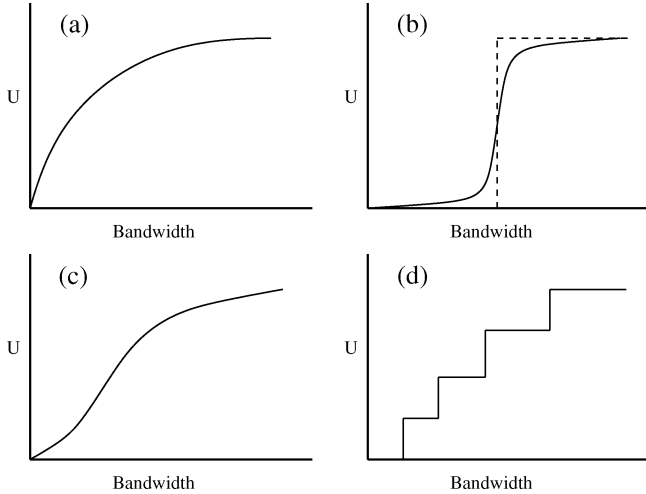
Fig. 1. Utility functions for different classes of applications. (a) Elastic. (b) Real-time. (c) Rate-adaptive. (d) Stepwise.

flow control algorithm that directly addresses the different QoS utility requirements.

III. APPLICATION-ORIENTED FLOW CONTROL PROBLEM

For a practical network application, the user may be concerned with the bandwidth allocation. However, a more important factor is the QoS performance which is measured by the application's utility function. The utility function of an application is a measurement of the application's QoS performance based on the provided network services, such as bandwidth, transmission delay and loss ratio [1]. As set out in the optimal flow control literature cited above, in this paper, we will assume that the utility is a function of the transmission bandwidth.

A. Utility Functions Within Network Applications

In [1], Shenker observes that traditional data applications, such as file transference, electronic mail and web browsing, are rather tolerant of throughput and time delays. This class of applications is called *elastic* traffic, and their utility functions can be described as a strictly concave function as shown in Fig. 1(a). Utility (performance) increases with an increase in bandwidth, but the marginal improvement is decreasing. This class of applications has been well studied in the recent optimal flow control literature.

Real-time applications, such as audio and video delivery, are generally delay sensitive and operate under strict QoS requirements. Unlike the *elastic* data traffic, these applications usually have an intrinsic bandwidth requirement because the data generation rate is independent of and isolated from network congestion. Thus, a degradation in bandwidth may result in serious packet drops and severe degradation in performance. A reasonable description of this class of utilities is close to a single step function as shown in Fig. 1(b) (solid line), which is convex but not concave at low bandwidths. Some hard real-time applications may require an exact step utility function as in Fig. 1(b) (dashed line).

There exists another class of real-time applications called rate-adaptive applications which have an ability to adjust their transmission rates in response to network congestion. Here the marginal utility increase in additional bandwidth is small for low and high bandwidths, and the utility curves may have a general shape as in Fig. 1(c).

There are some applications which may take a stepwise utility function as shown in Fig. 1(d). Such applications can be found in audio and video delivery systems via a layered encoding and transmission model [16], [26], [27]. For these applications, bandwidth allocations could be limited to distinct levels. The application's utility is increased only when an additional layer can be delivered owing to an increase in available bandwidth.

B. Utility-Based Resource Allocation Criteria

When considering different QoS requirements among network users, it may be undesirable to allocate bandwidth simply according to the conventional criteria such as max-min fairness [25] and proportional fairness [2]. Instead, a network should have the ability to allocate bandwidth resources to various users, addressing their real utility requirements. This has been the motivation for a new concept of utility max-min fairness suggested by Cao and Zegura [28].

Definition 2: A bandwidth allocation $x^* = [x_1^*, x_2^*, \ldots, x_S^*]^T$ is utility max-min fair, if it is feasible and for each user s, the utility $U_s(x_s^*)$ cannot be increased while still maintaining feasibility, without decreasing the utility $U_{s'}(x_{s'}^*)$ for some user s' with a lower utility $U_{s'}(x_{s'}^*) \leq U_s(x_s^*)$. Max-min fair allocation is recovered with

$$U_s(x_s) = x_s, \quad s = 1, \ldots, S.$$

Here we propose a new application-oriented fairness criterion: utility proportional fairness.

Definition 3: A bandwidth allocation $x^* = [x_1^*, x_2^*, \ldots, x_S^*]^T$ is utility proportionally fair, if it is feasible and for any other feasible allocation x,

$$\sum_{s \in \mathcal{S}} \frac{x_s - x_s^*}{U_s(x_s^*)} \leq 0. \tag{12}$$

As a special case, when $U_s(x_s) = x_s$, utility proportional fairness reduces to the well-known bandwidth proportional fairness criterion. The difference between utility proportional fairness and utility max-min fairness is analogous to the difference between (bandwidth) proportional fairness and (bandwidth) max-min fairness. In the next section, we will develop a new flow control algorithm to achieve utility fairness within a given network and study its properties in detail.

IV. UTILITY FAIR FLOW CONTROL ALGORITHM

Consider the flow control problem formulated in Section II-A. Each source s attains a non-negative and bounded utility $U_s(x_s)$ when it is allocated a bandwidth $x_s \in [m_s, M_s]$, where $m_s \geq 0$ and $M_s < \infty$ are the minimum and maximum transmission rates required by source s, respectively. The utility function $U_s(x_s)$ is assumed to be continuous and strictly

increasing in the interval $[m_s, M_s]$.[2] Without loss of generality, it can be assumed that $U_s(x_s) = 0$ when $x_s < m_s$ and $U_s(x_s) = U_s(M_s)$ when $x_s > M_S$.[3]

Next, we propose a new distributed algorithm that achieves utility proportional fairness in resource allocations, which has a close connection to the well-established optimal flow control algorithm set out in Section II-B, but also demonstrates significant differences.

A. A New Distributed Flow Control Algorithm

The new algorithm uses the same flow control structure as the optimal flow control approach [5] does, in which a link algorithm is deployed at each link to update the link price depending on the severity of link congestion, and a source algorithm is implemented at each source edge to adapt the transmission rate based on the feedback path price.

The link algorithm is the same as that of (9). At time $t + 1$, each link l updates its link price p_l according to

$$p_l(t+1) = \left[p_l(t) + \gamma \left(x^l(t) - c_l \right) \right]^+ \tag{13}$$

where $\gamma > 0$ is a small step size, and $x^l(t) = \sum_{s \in S_l} x_s(t)$ is the aggregate source rate at link l. Equation (13) says that if the aggregate source rate at link l exceeds the link capacity c_l, the link price will be increased; otherwise it will be decreased. The projection $[z]^+ = \max\{0, z\}$ ensures that the link price is always non-negative.

When each source s receives feedback congestion information, i.e., the path price, which is the sum of the link prices along its path, it adopts the following source algorithm to update the source rate:

$$x_s(t+1) = U_s^{-1} \left(\left[\frac{1}{p^s(t)} \right]_{U_s(m_s)}^{U_s(M_s)} \right) \tag{14}$$

where

$$p^s(t) = \sum_{l \in \mathcal{L}_s} p_l(t) \tag{15}$$

is the path price of source s, $[z]_a^b = \max\{a, \min\{b, z\}\}$, and U_s^{-1} is the inverse of U_s over the range $[U_s(m_s), U_s(M_s)]$. According to the definition of utility function, it is clear that $x_s(p^s)$ given in (14) is decreasing over the path price p^s. When $p^s \geq 1/U_s(m_s)$, source s is required to transmit at a minimum rate m_s. When $p^s \leq 1/U_s(M_s)$, source s transmits at a maximum rate M_s. In between, source s attains a utility factor of $1/p^s$.

Even though the source algorithms (14) and (8) both manage the transmission rate as a decreasing function of the path price, they are significantly different in their mathematical descriptions. We will show shortly that these two source algorithms may produce quite different resource allocation results.

[2]In this paper, we do not consider the layered audio/video delivery applications which have a discontinuous utility function as shown in Fig. 1(d) and the network flow control problem for such a class of applications will be dealt with in our future research.

[3]For matters of scalability, it can be further assumed that $0 \leq U_s(x_s) \leq 1$ and $U_s(M_s) = 1$.

B. Optimization and Convergence

The flow control algorithm (13) and (14) can be viewed as a distributed dual algorithm that solves the following optimization problem:

$$\max_{m_s \leq x_s \leq M_s} \quad \mathcal{U}(x) = \sum_{s=1}^{S} \mathcal{U}_s(x_s) \tag{16}$$

$$\text{subject to} \quad \sum_{s \in S_l} x_s \leq c_l, \quad l = 1, \dots, L \tag{17}$$

where

$$\mathcal{U}_s(x_s) = \int_{m_s}^{x_s} \frac{1}{U_s(y)} dy, \quad m_s \leq x_s \leq M_s \tag{18}$$

is a redefined "utility" function for source s.

The original utility function $U_s(x_s)$ is non-negative, continuous and strictly increasing over the range $x_s \in [m_s, M_s]$. Therefore, $\mathcal{U}_s(x_s)$ must be increasing and strictly concave. If the step size γ in (13) is selected to be appropriately small, the sequence (x, p) generated by the dual algorithm (13) and (14) will solve the maximization problem (16), (17).

Let

$$\bar{L} = \max_{s \in \mathcal{S}} |\mathcal{L}_s| \tag{19}$$

$$\bar{S} = \max_{l \in \mathcal{L}} |\mathcal{S}_l| \tag{20}$$

$$\alpha_1 = \max_{s \in \mathcal{S}} \max_{m_s \leq x_s \leq M_s} U_s(x_s)$$
$$= \max_{s \in \mathcal{S}} U_s(M_s) \tag{21}$$

$$\alpha_2 = \min_{s \in \mathcal{S}} \min_{m_s \leq x_s \leq M_s} U_s'(x_s) > 0 \tag{22}$$

where $|A|$ denotes the cardinality of the set A. In other words, $\bar{L}$ is the length of the longest path used by any given source, $\bar{S}$ is the number of the sources sharing the most congested link, α_1 is the upper bound on all the feasible utility $U_s(x_s)$, and α_2 is the lower bound on all $U_s'(x_s)$ over the range $[m_s, M_s]$.

We now state the main results regarding the convergence of the algorithm.

Theorem 1: Suppose the step size γ is selected to be

$$0 < \gamma < \frac{2\alpha_2}{\alpha_1^2 \bar{L} \bar{S}}$$

then the sequence $(x(t), p(t))$ generated by the flow control algorithm (13) and (14) will converge to a limit point (x^*, p^*), and x^* is the unique optimal solution for the maximization problem (16), (17).

Proof: The proof is given in the Appendix. $\square$

Remark 1: For each source s, we can also use the following Lagrangian first-order algorithm:

$$x_s(t+1) = \left[x_s(t) + \gamma \left(\frac{1}{U_s(x_s(t))} - p^s(t) \right) \right]_{m_s}^{M_s} \tag{23}$$

to update source rate x_s instead of using the dual algorithm (14). According to our earlier results shown in [29], this gentle and smooth adaptation gives a better performance than that of the dual algorithm (14), especially when the path price cannot be fed

back perfectly to its source due to network implementation limitations, i.e., a random exponential marking (REM) [23] strategy must then be deployed.

C. Utility Proportional Fairness

When the flow control algorithm (13) and (14) converges to the equilibrium (x^*, p^*), the objective function (16) is maximized within the feasible solution. For all feasible allocation $x \neq x^*$, the optimality condition is

$$\sum_{s \in \mathcal{S}} \frac{\partial \mathcal{U}_s(x_s^*)}{\partial x_s}(x_s - x_s^*) = \sum_{s \in \mathcal{S}} \frac{x_s - x_s^*}{U_s(x_s^*)} < 0 \qquad (24)$$

where the strict inequality follows from the strict concavity of $\mathcal{U}_s(x_s)$. According to Definition 3, it is clear that, at optimality, the resource allocation x^* is utility proportionally fair. Moreover, if source s transmits at a nontrivial rate $m_s < x_s^* < M_s$, then it achieves a utility factor $U_s(x_s^*) = 1/p^{s^*}$, which is reciprocal to the sum of the link prices along its network path.

D. Utility Max-Min Fairness

The original utility max-min fair flow control algorithm given by Cao and Zegura [28] is not distributive and each link must know the utility functions from all other sources that traverse such a link. Here we give a new distributive algorithm to achieve the same objective.

For each source s, if the path price is redefined as

$$p^s(t) = \max_{l \in \mathcal{L}_s} p_l(t) \qquad (25)$$

which is the maximum of the link prices other than the sum of the link prices along the path, then flow control algorithms (13) and (14) will provide a utility max-min fair allocation within the network. In this schema, the link price can be viewed as the lowest packet priority that each link is willing to support for the arrival packets in a priority service network. Meanwhile, each source must assign to its packets the highest priority requirement (25) along its path to avoid packets being discarded at congested links. To achieve (utility) max-min fairness, an *explicit* information feedback mechanism must be available within the network to inform the end user of the maximum link price (highest priority requirement) along the path. However, for (utility) proportional fairness to occur, the path price (sum of the link prices) can be *implicitly* measured by the packet queuing delay in the current Internet environment.

E. Utility Fairness Versus Utility Maximization

Even though utility proportional fairness and utility max-min fairness can provide a fair solution for resource allocation in existing networks, the result may produce a lower aggregate utility than that seen in the traditional optimal flow control strategies.

Consider the following example in which two flows share a link with 3 units of bandwidth. Their utility functions are both

$$U(x) = \begin{cases} x/20 & \text{if } 0 \leq x \leq 2 \\ 9x - 17.9 & \text{if } 2 < x \leq 2.1 \\ 1 & \text{if } x \geq 2.1. \end{cases}$$

This utility function could be a real-time application that has a hard bandwidth requirement of 2.1 units. The utility fair allocation given in this paper assigns 1.5 units of bandwidth to each flow, and thus the utility of each flow will be 0.075, i.e., no flow is usable. Alternatively, if we were to maximize the total utility in the system, then we would have allocated 2.1 units to one of the flows, so that at least this flow would be able to meet the required bandwidth and achieve a utility of 1. This example shows that favoring utility fairness could in practice lead to a less meaningful allocation as opposed to maximizing the total system utility.

The final solution to this practical problem is to build high speed networks which will have sufficient bandwidth to support various real-time applications. However, as long as the bandwidth is insufficient, an *admission control* policy is essential for traffic management if the network cannot support all admission requirements.

V. EXAMPLES AND SIMULATIONS

In this section, we present the numerical results for two examples. In the first example, we show the difference in resource allocations between the utility proportionally fair strategy and the conventional optimal flow control approach. In the second example, we apply the algorithms proposed in this paper to a specific network to demonstrate their dynamic behaviors.

A. Example 1

Consider a single link with a capacity of c shared by two sources. Source 1 attains a utility of $\log(x_1 + 1)$ and source 2 attains a utility of $2\log(x_2 + 1)$.

Using a simple calculation, it can be shown that:
1) in the utility proportionally fair approach

$$x_1 = c - \frac{\sqrt{4c+9} - 3}{2} \qquad (26)$$

$$x_2 = \frac{\sqrt{4c+9} - 3}{2}; \qquad (27)$$

2) in the utility maximization-based optimal flow control approach

$$x_1 = \begin{cases} 0 & \text{if } c \leq 1 \\ \frac{c-1}{3} & \text{if } c > 1 \end{cases} \qquad (28)$$

$$x_2 = \begin{cases} c & \text{if } c \leq 1 \\ \frac{2c+1}{3} & \text{if } c > 1. \end{cases} \qquad (29)$$

Fig. 2(a) shows the bandwidth allocation for these two strategies when link capacity c varies from 0 to 10. Their associated utilities are given in Fig. 2(b). In the utility proportionally fair (UPF) approach, both sources attain the same utility. In the optimal flow control (OFC) approach, the total utilities are maximized. However, in this situation source 1 is poorly treated, and source 2 is favored in the extreme, even though it may only need a smaller bandwidth to achieve the same performance level. In particular, when link bandwidth is scarce, i.e., $0 < c \leq 1$, in the OFC approach, source 1 is totally prohibited from transmission as all of the bandwidth is allocated to source 2. It is clear that the utility maximization derived OFC approach can lead to a seriously unfair environment in network resource allocations.
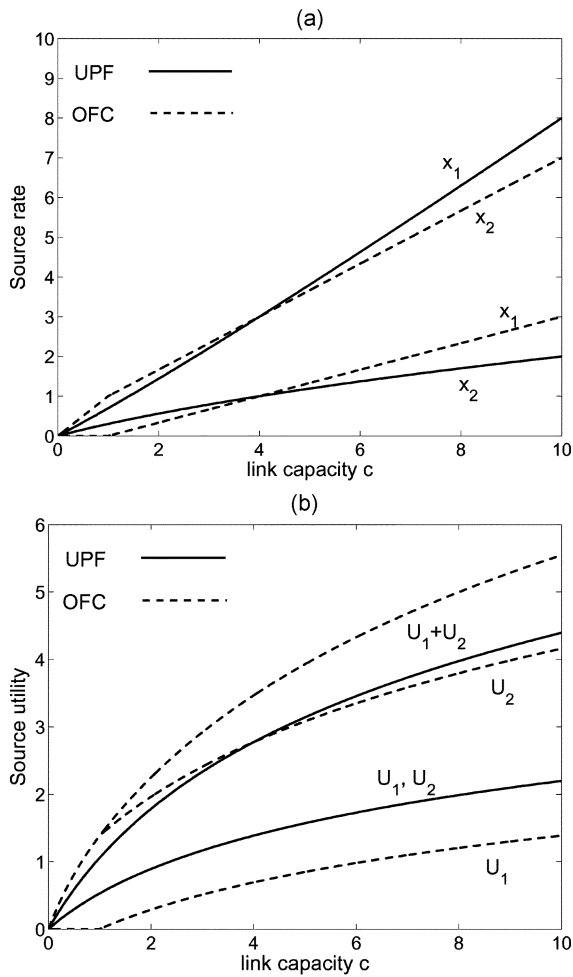
Fig. 2. Comparison results of Example 1: utility proportionally fair (UPF) versus optimal flow control (OFC). (a) Difference of bandwidth allocation between UPF and OFC. (b) Difference of source utility between UPF and OFC.
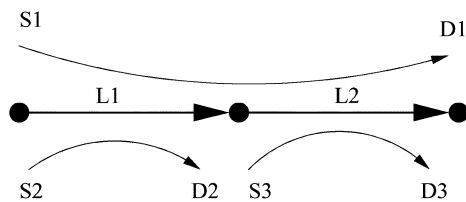

Fig. 3. Network topology of Example 2.

B. Example 2

Consider the following simple network, as shown in Fig. 3, which consists of two links L1 and L2 with a capacity of 10 Mb/s and shared by three sources S1, S2, and S3. S1 traverses link L1 and L2, and S2 and S3 traverse L1 and L2, respectively.

Their utilities are shown in Fig. 4(a), in which $U_1(x_1) = 1/(1+e^{-2(x_1-5)})$, $U_2(x_2) = 0.125x_2$, and $U_3(x_3) = \log(x_3+1)/\log 11$. The three sources have their maximum rate requirement set at 10, 8, and 10 Mb/s, respectively.

In the simulation, each link and source update their algorithm (13) and (14) iteratively every 50 ms, at a step size $\gamma = 0.05$. The simulation consists of two stages.
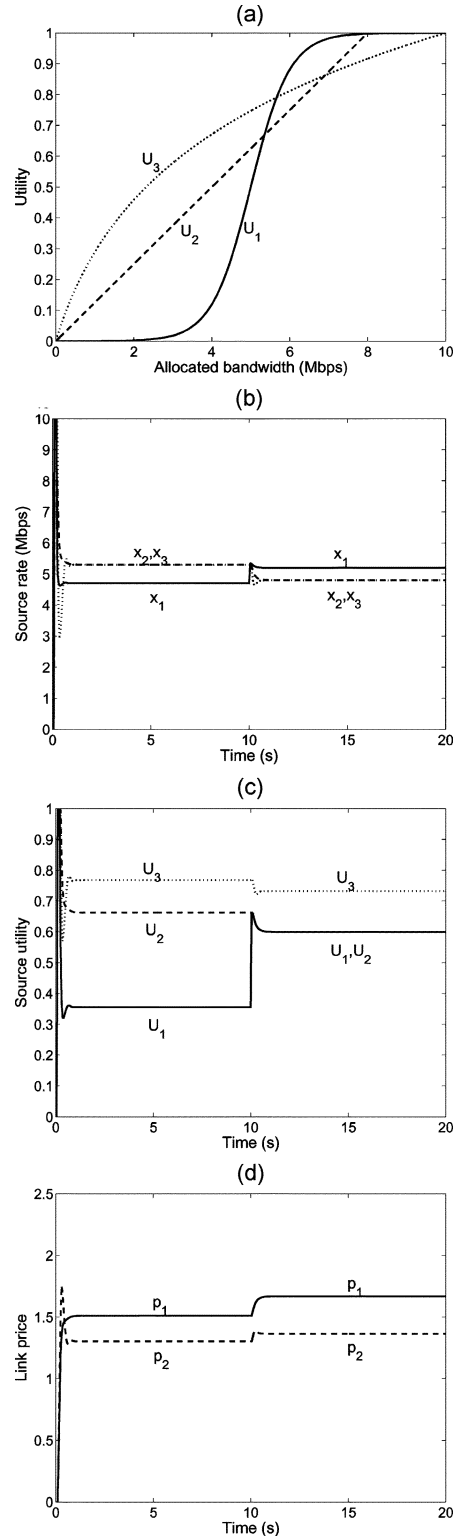

Fig. 4. Simulation results of Example 2. (a) Source utility functions for S1, S2, and S3. (b) Convergence of source rates. (c) Associated source utilities. (d) Corresponding link prices.

- Stage 1: $t = 0 \rightarrow 10$ s, utility proportional fairness is used and the path price (15) is defined as the sum of the link prices.
- Stage 2: $t = 10 \rightarrow 20$ s, utility max-min fairness is deployed and the path price (25) is defined as the maximum of the link prices.

The simulation results are given in Fig. 4(b)–(d).

- Stage 1: In utility proportional fairness, source rates $x_1(t)$, $x_2(t)$, and $x_3(t)$ converge in equilibrium (4.7024, 5.2976, 5.2976), utilities U_1, U_2, and U_3 converge at (0.3555, 0.6622, 0.7674), and link prices $p_1(t)$ and $p_2(t)$ converge at (1.5101, 1.3031). It can be verified that at convergence, for all sources, $U_s p^s = 1$. S1 achieves a lower utility since it traverses two congested links.
- Stage 2: In utility max-min fairness, source rates converge at (5.2022, 4.7978, 4.7978), S1 and S2 achieve the same utility rate of 0.5997 due a bottle neck link price 1.6674 at L1. S3 achieves a higher utility rate of 0.7329 and the link price of L2 is 1.3644.

This confirms that our flow control algorithms are efficient and able to provide either utility-based proportional fairness or utility-based max-min fairness in resource allocations among competing applications. Moreover, the utility function, i.e., $U_1(x_1)$ and $U_2(x_2)$, does not need to satisfy the critical strict concavity condition which is required in the standard optimal flow control approach.

VI. CONCLUSION

In this paper, we have examined the limitations of the widely studied optimal flow control strategy for computer networks, and we have shown that the utility maximization approach could lead to a seriously unfair situation in resource allocations. Therefore, we propose a new resource allocation criterion named "utility proportional fairness," along with a new distributed flow control algorithm that we have developed to achieve utility fairness in a given network. This new algorithm shares the same link pricing policy with the standard optimal flow control strategy, but adopts a different rate adjustment in the source algorithm. When each source receives information of the path price which constitutes the sum of the link prices in its path, it sends traffic at a rate for which the attained utility is equal to the reciprocal of the path price. We have shown that at convergence, the bandwidth is properly allocated, and the utility achieved by each source comes into proportional fairness. Furthermore, if the path price is defined as the maximum, other than the sum of the link prices along the path, utility max-min fairness is guaranteed within the same flow control framework.

The algorithm presented in this paper requires only that the source utility functions be positive, strictly increasing and bounded over the available bandwidth. It does not require the strict concavity condition on the utilities that is strongly desired by the standard optimal flow control approach. Therefore, our algorithm is more suitable in providing efficient flow control and fair resource allocation for real-time service networks.

While this paper only addresses the fair flow control problem for single-path networks, the results we have obtained can be easily extended to multiple-path and multirate multicast networks. Since the algorithm uses the same link algorithm found in the standard optimal flow control approach, the techniques developed for optimal flow control, such as RED [22], REM [23], and path pricing via transmission delay [15], can be used here to enhance the performance of our proposed algorithm.

APPENDIX
PROOF OF THEOREM 1

In this Appendix, we will consistently use vector notations. For a vector $z = (z_1, z_2, \ldots, z_n)^{\mathrm{T}}$, $\|z\|_2$ denotes the Euclidean norm, $\|z\|_1 = \sum_i |z_i|$, $\|z\|_\infty = max_i |z_i|$, and $\|z\|$ without a subscript denotes any norm. For a matrix z, $\|z\|$ denotes the correspondingly induced norm.

It will sometimes be convenient to represent the information of $\mathcal{L}_s$ and $\mathcal{S}_l$ (see our definition in Section II-A) in terms of a $L \times S$ routing matrix R in which $R_{l,s} = 1$ if link l is traversed by source s, and 0 otherwise. The link capacity constraints (3) and (17) can be written as $Rx \le c$, where $c = [c_1, c_2, \ldots, c_L]^{\mathrm{T}}$.

According to the Lagrangian formulation (6), the dual function of the optimization problem (16), (17) is defined as follows:

$$D(p) = \max_{m_s \le x_s \le M_s} \left(\sum_{s=1}^{S} [\mathcal{U}_s(x_s) - x_s p^s] + \sum_{l=1}^{L} p_l c_l \right) \quad (30)$$

$$= \sum_{s=1}^{S} [\mathcal{U}_s(x_s(p^s)) - x_s p^s] + \sum_{l=1}^{L} p_l c_l \quad (31)$$

where

$$p^s = \sum_{l \in \mathcal{L}_s} p_l \quad (32)$$

and

$$x_s(p) = x_s(p^s) = U_s^{-1} \left(\left[\frac{1}{p^s} \right]_{U_s(m_s)}^{U_s(M_s)} \right) \quad (33)$$

is given by the source algorithm (14) that solves the maximization problem in (30).

The link algorithm (13) can be further viewed as a gradient projection algorithm that solves a dual problem

$$\mathbf{D}: \quad \min_{p \ge 0} D(p). \quad (34)$$

The following proof is closely related to the proof of Theorem 1 in [5].

For the property of the dual function $D(p)$, we have the following lemma directly from the definition of $\mathcal{U}_s(x_s)$.

Lemma 1: The dual function $D(p)$ is convex, lower bounded and continuously differentiable.

For any given price vector $p \ge 0$, we define $\beta_s(p)$ by

$$\beta_s(p) = \begin{cases} \dfrac{U_s^2(x_s(p))}{U_s'(x_s(p))}, & \text{if } \dfrac{1}{U_s(M_s)} \le p^s \le \dfrac{1}{U_s(m_s)} \\ 0, & \text{otherwise} \end{cases}$$

where $x(p)$ is given in (33). We will use $x_s(\cdot)$ both as a function of the (scalar) path price p^s and of the vector price p.

Let $B(p)$ be the $S \times S$ diagonal matrix defined by

$$B(p) = Diag\left(\beta_s(p), s \in \mathcal{S}\right). \quad (35)$$

Recall the routing matrix R defined in the beginning of the Appendix.

Lemma 2: The Hessian of D is given by $\nabla^2 D(p) = RB(p)R^{\mathrm{T}}$, where it exists.

Proof: Let $(\partial x/\partial p)(p)$ denote the $S \times L$ Jacobian matrix whose (s,l) element is $(\partial x_s/\partial p_l)(p)$. According to (33)

$$\frac{\partial x_s}{\partial p_l}(p) = \begin{cases} -\frac{U_s^2(x_s(p))}{U_s'(x_s(p))}R_{l,s}, & \text{if } \frac{1}{U_s(M_s)} \leq p^s \leq \frac{1}{U_s(m_s)} \\ 0, & \text{otherwise.} \end{cases}$$

Using (35), we have

$$\frac{\partial x}{\partial p}(p) = -B(p)R^{\mathrm{T}}. \tag{36}$$

Thus, from (31), we have $\nabla D(p) = c - Rx(p)$ and, hence

$$\nabla^2 D(p) = -R\left(\frac{\partial x}{\partial p}(p)\right) = RB(p)R^{\mathrm{T}}. \tag{37}$$

$\square$

Recall $\bar{L}$, $\bar{S}$, α_1 and α_2 defined in (19)–(22), and we have:
Lemma 3: $\nabla D(p)$ is Lipschitz with

$$\|\nabla D(q) - \nabla D(p)\|_2 \leq \frac{\alpha_1^2 \bar{L}\bar{S}}{\alpha_2}\|q - p\|_2 \tag{38}$$

for any vector $p, q \geq 0$.

Proof: Using Lemma 2, we will show that $\|\nabla^2 D(p)\|_2 = \|RB(p)R^{\mathrm{T}}\| \leq (\alpha_1^2 \bar{L}\bar{S}/\alpha_2)$. The lemma then follows from [30, Theorem 9.19].[4]

With the definition $B(p)$ in (35)

$$\|\nabla^2 D(p)\|_2 = \|RB(p)R^{\mathrm{T}}\|_2 \tag{39}$$

$$\leq \frac{\alpha_1^2}{\alpha_2}\|RR^{\mathrm{T}}\|_2. \tag{40}$$

Since (see [31, p. 635])

$$\|RR^{\mathrm{T}}\|_2^2 \leq \|RR^{\mathrm{T}}\|_\infty \|RR^{\mathrm{T}}\|_1$$

and RR^{T} is symmetric, $\|RR^{\mathrm{T}}\|_\infty = \|RR^{\mathrm{T}}\|_1$, hence

$$\begin{aligned}
\|RR^{\mathrm{T}}\|_2 &\leq \|RR^{\mathrm{T}}\|_\infty \\
&= \max_l \sum_{l'} [RR^{\mathrm{T}}]_{ll'} \\
&= \max_l \sum_{l'} \sum_s R_{l,s} R_{l',s} \\
&\leq \bar{L} \max_l \sum_s R_{l,s} \\
&\leq \bar{L}\bar{S}.
\end{aligned}$$

Together with (40), we have (38). which is desirable. $\square$

Since $x(p)$ in (33) is continuous, the dual function $D(p)$ is lower bounded from Lemma 1 and $\nabla D(p)$ is $(\alpha_1^2 \bar{L}\bar{S}/\alpha_2)$ Lipschitz from Lemma 3. Let $0 < \gamma < (2\alpha_2/\alpha_1^2 \bar{L}\bar{S})$, any sequence $p(t)$ generated by the gradient projection algorithm (13) converges to a limit point p^*, which is the optimal solution for the dual problem **D**. Meanwhile, $x^* = x(p^*)$ is the unique solution for the primal problem (16), (17) (see [31, p. 214]). Thus, the proof of Theorem 1 is complete.

---

[4]This was pointed out by Edward Fan at UCLA in March 2002 and corrects an error in the original proof in [5].

REFERENCES

[1] S. Shenker, "Fundamental design issues for the future Internet," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1176–1188, Sep. 1995.
[2] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, Jan. 1997.
[3] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, Mar. 1998.
[4] R. J. Gibbens and F. P. Kelly, "Resource pricing and the evolution of congestion control," *Automatica*, vol. 35, pp. 1969–1985, Dec. 1999.
[5] S. H. Low and D. E. Lapsley, "Optimal flow control, I: basic algorithm and convergence," *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, Dec. 1999.
[6] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000.
[7] H. Yaiche, R. R. Mazumber, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 667–678, Oct. 2000.
[8] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Contr. Syst. Mag.*, vol. 22, no. 1, pp. 28–43, Feb. 2002.
[9] R. La and V. Anantharam, "Utility-based rate control in the Internet for elastic traffic," *IEEE/ACM Trans. Netw.*, vol. 10, no. 2, pp. 272–286, Apr. 2002.
[10] E. Altman, T. Basar, and R. Srikant, "Nash equilibria for combined flow control and routing in networks: asymptotic behavior for a large number of users," *IEEE Trans. Autom. Contr.*, vol. 47, no. 6, pp. 917–930, Jun. 2002.
[11] T. Alpcan and T. Basar, "Distributed algorithms for Nash equilibria of flow control games," in *Annals of Dynamic Games*, A. Nowak, Ed. Cambridge, MA: Birkhauser, 2003.
[12] ——, "A utility-based congestion control scheme for Internet-style networks with delay," in *Proc. IEEE INFOCOM 2003*, pp. 2039–2048.
[13] F. P. Kelly, "Fairness and stability of end-to-end congestion control," *Eur. J. Contr.*, vol. 9, pp. 149–165, 2003.
[14] K. Kar, S. Sarkar, and L. Tassiulas, "Optimization-based rate control for multipath sessions," Inst. Syst. Res., Univ. Maryland, Baltimore, MD, Tech. Rep. 2001-1, 2001.
[15] W. H. Wang, M. Palaniswami, and S. H. Low, "Optimal flow control and routing in multiple paths networks," *Perform. Eval.*, vol. 52, no. 2–3, pp. 119–132, 2003.
[16] K. Kar, S. Sarkar, and L. Tassiulas, "A scalable low-overhead rate control algorithm for multirate multicast sessions," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 8, pp. 1541–1557, Oct. 2002.
[17] W. H. Wang, M. Palaniswami, and S. H. Low, "Necessary and sufficient conditions of optimal flow control in multirate multicast networks," *Proc. Inst. Electr. Eng. —Commun.*, vol. 150, pp. 385–390, Oct. 2003.
[18] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[19] P. Marbach, "Priority service and max-min fairness," *IEEE/ACM Trans. Netw.*, vol. 11, no. 10, pp. 733–746, Oct. 2003.
[20] V. Jacobson and M. J. Karels, "Congestion avoidance and control," in *Proc. ACM SIGCOMM'88*, Aug. 1988, pp. 314–329.
[21] L. S. Brakmo and L. L. Peterson, "TCP Vegas: end to end congestion avoidance on a global Internet," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 8, pp. 1465–1480, Oct. 1995.
[22] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. Netw.*, vol. 1, no. 4, pp. 397–413, Aug. 1993.
[23] S. Athuraliya, S. H. Low, V. H. Li, and Q. Yin, "REM: active queue management," *IEEE Network*, vol. 15, no. 3, pp. 48–53, May–Jun. 2001.
[24] S. H. Low, "A duality model of TCP and queue management algorithms," *IEEE/ACM Trans. Netw.*, vol. 11, no. 4, pp. 525–536, Aug. 2003.
[25] D. Bertsekas and R. Gallager, *Data Networks*, 2 ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[26] T. Bially, B. Gold, and S. Seneff, "A technique for adaptive voice flow control in integrated packet networks," *IEEE Trans. Commun.*, vol. 28, no. 3, pp. 325–333, Mar. 1980.
[27] F. Kishino, K. Manabe, Y. Hayashi, and H. Yasuda, "Variable bit-rate coding of video signals for ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 5, pp. 801–806, Jun. 1989.

[28] Z. Cao and E. W. Zegura, "Utility max-min: an application-oriented bandwidth allocation scheme," in *Proc. IEEE INFOCOM'99*, pp. 793–801.

[29] W. H. Wang, "Optimal flow control in packet switched networks," Ph.D. dissertation, Univ. Melbourne, Victoria, Australia, Apr. 2003.

[30] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hall, 1976.

[31] D. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

Wei-Hua Wang received the B.E. degree from Xi'an Jiaotong University, China, in 1992, the M.E. degrees from Northeastern University, China, in 1995 and from Nanyang Technological University, Singapore, in 1997, and the Ph.D. degree from the University of Melbourne, Victoria, Australia, in 2004, all in electrical engineering.

He was a Lecturer at the Department of Electronic Engineering, Fudan University, China, from 1997 to 1999, and is currently a Research Fellow in the Department of Electrical and Electronic Engineering, University of Melbourne. His research interests include network flow control, real-time Internet services, optimization, nonlinear systems and robust control theory.

Marimuthu Palaniswami (SM'95) received the B.E. (Hons.) degree from the University of Madras, Chennai, India, the M.E. degree from the Indian Institute of Science, Bangalore, India, the M.Eng.Sc. degree from the University of Melbourne, Parkville, Victoria, Australia, and the Ph.D. degree from the University of Newcastle, Newcastle, Australia.

He has been with the University of Melbourne for over 16 years. He has published more than 180 refereed papers with a majority of them in prestigious IEEE journals and conferences. His research interests are in the fields of computer networks, sensors and sensor networks, machine learning, neural network, pattern recognition, signal processing and control. He is the convener for Australian Research Network on Sensor Network. He is the Co-Director of Centre of Expertise on Networked Decision and Sensor Systems.

Dr. Palaniswami received a Foreign Specialist Award from the Ministry of Education, Japan, in recognition of his contributions to the field of machine learning. He has served as associate editor for journals/transactions including IEEE TRANSACTIONS ON NEURAL NETWORKS and *Computational Intelligence for Finance*. He is the associate editor for the *International Journal of Computational Intelligence and Applications* and serves on the editorial board of the *ANZ Journal on Intelligent Information Processing Systems*. He is also the Subject Editor for the *International Journal on Distributed Sensor Networks*.
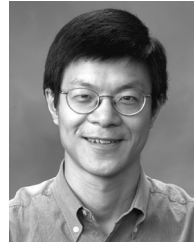
Steven H. Low (M'92–SM'99) received the B.S. degree from Cornell University, Ithaca, NY, and the Ph.D. degree from the University of California at Berkeley.

He is a Professor of the Computer Science and Electrical Engineering Departments at Caltech, Pasadena.

Dr. Low was a co-recipient of the IEEE William R. Bennett Prize Paper Award in 1997 and the 1996 R&D 100 Award. He was on the editorial board of IEEE/ACM TRANSACTIONS ON NETWORKING from 1997 to 2006 and of *Computer Networks Journal* from 2003 to 2005. He is on the editorial boards of *ACM Computing Surveys*, *NOW Foundations* and *Trends in Networking*. He is a Senior Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and a Co-Editor of the Springer Book Series on Optimization and Control of Communication Systems: Theory and Applications. He is a member of the Networking and Information Technology Technical Advisory Group for the U.S. President's Council of Advisors on Science and Technology (PCAST).