Achieving Network Optima Using Stackelberg Routing Strategies

Yannis A. Korilis, *Member, IEEE*, Aurel A. Lazar, *Fellow, IEEE*, and Ariel Orda, *Member, IEEE*

***Abstract*—In noncooperative networks users make control decisions that optimize their individual performance objectives. Nash equilibria characterize the operating points of such networks. Nash equilibria are generically inefficient and exhibit suboptimal network performance. Focusing on routing, a methodology is devised for overcoming this deficiency, through the intervention of the network manager. The manager controls part of the network flow, is aware of the noncooperative behavior of the users and performs its routing aiming at improving the overall system performance. The existence of *maximally efficient* strategies for the manager, i.e., strategies that drive the system into the global network optimum, is investigated. A maximally efficient strategy of the manager not only optimizes the overall performance of the network, but also induces an operating point that is efficient with respect to the performance of the individual users (Pareto efficiency). Necessary and sufficient conditions for the existence of a maximally efficient strategy are derived, and it is shown that they are met in many cases of practical interest. The maximally efficient strategy is shown to be unique and it is specified explicitly.**

***Index Terms*— Nash equilibria, networking games, network management, routing.**

NOMENCLATURE

$\mathcal{L} = \{1, \ldots, L\}$	Set of links.
$\mathcal{L}^i = \{1, \ldots, L^i\}$	Set of links receiving flow from user i.
$\mathcal{I} = \{1, \ldots, I\}$	Set of self-optimizing users.
$\mathcal{I}_0 = \mathcal{I} \cup \{0\}$	Set of all users.
$\mathcal{I}_l = \{1, \ldots, I_l\}$	Set of users sending flow on link l.
c_l	Capacity of link l.
$\mathbf{c} = (c_1, \ldots, c_L)$	Capacity configuration.
$C = \sum_{l \in \mathcal{L}} c_l$	Total network capacity.
$c_l^i = c_l - \sum_{j \in \mathcal{I}_0 \setminus \{i\}} f_l^j$	Capacity of link l seen by user i.
r^i	Throughput demand of user i.
$r = \sum_{i \in \mathcal{I}} r^i$	Total demand of the followers.
$R = r + r^0$	Total follower and manager demand.
f_l^i	Flow of user i on link l.
$\mathbf{f}^i = (f_1^i, \ldots, f_L^i)$	Routing strategy of user i.
$\mathbf{f} = (\mathbf{f}^0, \mathbf{f}^1, \ldots, \mathbf{f}^I)$	Routing strategy profile.
$\mathbf{f}^{-i}$	Strategy profile of all users except the ith.
$f_l = \sum_{i \in \mathcal{I}_0} f_l^i$	Total flow on link l.
$J^i(\mathbf{f})$	Cost of user i under profile $\mathbf{f}$.
$J(\mathbf{f}) = \sum_{i \in \mathcal{I}_0} J^i(\mathbf{f})$	Total cost under profile $\mathbf{f}$.
$(f_1^*, \ldots, f_L^*)$	Network optimum.
J^*	Total cost at the network optimum.
λ^*	Lagrange multiplier for the network optimum.
$\mathcal{N}^0(\mathbf{f}^0)$	Nash equilibrium of followers under manager strategy $\mathbf{f}^0$.
$\underline{r}^0$	Leader threshold.

Manuscript received June 14, 1995; revised May 16, 1996; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor U. Shankar. This work was supported by the Office of Naval Research under Contract N00014-90-J-1289.

Y. A. Korilis is with Bell Labs/Lucent Technologies, Holmdel, NJ 07733 USA.

A. A. Lazar is with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA.

A. Orda is with the Department of Electrical Engineering and the S. Neaman Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel.

Publisher Item Identifier S 1063-6692(97)01636-1.

I. INTRODUCTION

CONTROL DECISIONS in large scale networks are often made by each user independently, according to its own individual performance objectives.[1] Such networks are henceforth called *noncooperative*, and game theory [1], [2] provides the systematic framework to study and understand their behavior. The operating points of a noncooperative network are the *Nash equilibria* of the underlying game, that is, the points where unilateral deviation does not help any user to improve its performance. Game theoretic models have been employed in the context of flow control [3]–[6], routing [7], [8] and virtual path bandwidth allocation [9] in modern networking. These studies mainly investigate the structure of the Nash equilibria and provide valuable insight into the nature of networking under decentralized and noncooperative control.

Nash equilibria are generically inefficient [10] and exhibit suboptimal network performance. This deficiency can be overcome with the intervention of a network agent, namely the network designer or manager, that architects the network so that the resulting equilibria are efficient according to some systemwide criterion. In essence, the designer/manager architects the Nash equilibria by setting the rules of the networking game. Various methods have been proposed for architecting Nash equilibria.

1) Through pricing mechanisms. In [11], [12], pricing strategies that lead to efficient usage of network resources are investigated.

[1] The term "user" may refer to a network user itself or, in case that the user's traffic consists of multiple connections, to individual connections that are controlled independently.

2) By regulating service disciplines. In [13], it is shown that a proper queue scheduling discipline can guarantee an equilibrium point with desirable properties.
3) Through proper network design. In [14], it is shown that, by making appropriate topology design and capacity allocation decisions, the network designer can choose a systemwide efficient equilibrium.

The above approaches demand either the addition of a new component to the networking structure, such as prices, or else *a priori* design decisions on the resource configuration and/or the service disciplines of the network. In the present study, we propose a method for architecting noncooperative equilibria in the *run time phase*, i.e., during the actual operation of the network. This approach is based on the observation that, apart from the flow generated by the self-optimizing users, typically, there is also some network flow that is controlled by a central entity, that will be referred to as the "manager." Typical examples are the traffic generated by signaling and/or control mechanisms, as well as traffic of users that belong to virtual networks. The manager attempts to optimize the system performance, through the control of its portion of the flow.

The role of the manager in a noncooperative network is investigated using routing as a control paradigm. The network is shared by a set of noncooperative users, each shipping its flow in a way that optimizes its individual performance objective. The noncooperative routing scenario applies to various modern networking environments. The Internet Protocol (both IPv4 and the current IPv6 Specification), for example, provides the option of source routing [15], [16], that enables the user to determine the path(s) its flow follows from the source to the destination. Another example is the flexible routing service as specified in the Q.1211 CCITT Recommendation for the standardized capability set of Intelligent Networks (IN CS-1) [17]. One of the goals of this service is to route calls over particular facilities based on the subscriber's routing preference list or distribution algorithm. Flexible routing was one of the services that were successfully implemented in Ameritech's AIN 0.0 technical trial, in April 1992 [18].

The manager has the following goals and capabilities: 1) it aims to optimize the overall network performance according to some systemwide efficiency criterion and 2) it is cognizant of the noncooperative behavior of the users and performs its routing based on this information. The first property makes the manager just another user, whose performance objective coincides with that of the network. The second property, however, enables the manager to predict the response of the noncooperative users to any routing strategy that it chooses, and hence determine a strategy that would pilot them to an operating point that optimizes the overall network performance. Instead of *reacting* to the routing strategies of the users, the manager *fixes* this optimal strategy and lets the users converge to their respective equilibrium. This is a typical scenario of a *Stackelberg game* [1], where the manager acts as a *leader*, that imposes its strategy on the self-optimizing users that behave as *followers*.[2] Stackelberg strategies have been investigated in the context of flow control in [19], and routing in [20]. In these references, however, the leader was a selfish user concerned about its own rather than the system's performance.

As an application of the proposed management scheme, consider, for example, a metropolitan-area network (MAN), where various institutions and organizations route their flows independently, according to certain performance criteria. An organization, however, may choose to request from the MAN provider a set of guaranteed and preassigned virtual paths (VP's) for the use of its own customers, building this way a virtual private network (VPN). While individual user flows might be controlled independently within the VPN, it is the task of the network provider to determine the routing of the VP's that compose the VPN. The network provider can, then, act as "the manager," and implement the VPN in a way that optimizes the overall performance of the network.

We investigate the existence of routing strategies of the manager that drive the system to the network optimum, i.e., to the point that corresponds to the solution of a routing problem, in which the manager has full control over the *entire* flow offered to the network. Such a strategy is called a *maximally efficient* strategy of the manager. As will be explained in the sequel, a maximally efficient strategy of the manager not only optimizes the overall performance of the network, but also drives the system to an operating point that is efficient from the perspective of the individual users, in the sense that there is no other point that improves the grade-of-service (GoS) received by a user without degrading the performance of some other user (Pareto efficiency). Intuitively, one would expect that the manager cannot enforce the network optimum, since it controls only part of the flow, while the rest is controlled by noncooperative users. Surprisingly, this study shows that in many cases the manager does have this capability.

The methodology will be developed for a simple network consisting of a common source and a common destination node interconnected by a number of parallel links. Systems of parallel links, albeit simple, represent an appropriate model for seemingly unrelated networking problems. Consider, for example, a network in which resources are preallocated to various routing paths that do not interfere. Such scenaria are common in modern networking. In broadband networks, for instance, bandwidth is separated among different virtual paths, resulting effectively in a system of parallel and noninterfering "links" between source/destination pairs. Moreover, to reduce the complexity of routing mechanisms, the network might present the users with a limited set of paths between source and destination, hiding the underlying physical topology. Another example is that of a corporation or organization that receives service from a number of different network providers. The corporation can split its total flow over the various network facilities (according to performance and cost considerations), each of which can be represented as a "link" in the parallel link model. Finally, it should be noted that routing, as a control paradigm, applies not only to the allocation of paths to messages and connections in communication networks, but in fact to any problem of splitting load among several resources, e.g., distribution of tasks among multiple processors. Consider, for example, a multimedia network with several servers that

[2] The terms "manager" and "leader," as well as "users" and "followers," will be used interchangeably.

are shared by the network customers: each customer distributes its applications among the servers, while competing with the other customers on the common available resources, resulting in effect with a routing game. Modeling each resource (e.g., multimedia server) as a "link," the parallel links model considered in our study fits well such scenaria.

We derive a *necessary and sufficient* condition that guarantees that the manager can enforce an equilibrium that coincides with the network optimum. The condition requires that the flow controlled by the manager exceeds a certain threshold. When this condition is satisfied, we show that the maximally efficient strategy of the manager is unique and we specify its structure explicitly. Finally, we investigate the dependency of the manager's threshold on the number of the users and their traffic characteristics.

An important question is then in place: In practice, does the manager control enough flow to meet the required threshold? Our analysis indicates that under moderate and heavy load conditions, the manager's threshold is small compared to the total flow of the self-optimizing users.[3] Therefore, in most practical cases of interest (moderate/heavy loading conditions) the manager is able to achieve, through limited control, the same system performance as in the case of centralized control. When, on the other hand, the manager does not have enough traffic to enforce the network optimum, our analysis provides guidelines for actions that it can take in order to meet the required threshold.

The outline of the paper is the following. In Section II, we present the parallel links model and formulate the problem. Section III gives an outline of the main results. The structure of the network optimum and Nash equilibrium are briefly described in Section IV. In Section V, we investigate the simplest Stackelberg routing game, where, except for the manager, there is a single self-optimizing user. The general multifollower Stackelberg routing game is presented in Section VI. Some practical issues are discussed in Section VII, while Section VIII summarizes the results and delineates their implications.

II. Model and Problem Formulation

We consider a set $\mathcal{I} = \{1, \cdots, I\}$ of users, that share a set $\mathcal{L} = \{1, \cdots, L\}$ of communication links, interconnecting a common source to a common destination node. The users are noncooperative, in the sense that each user routes its flow in a way that optimizes its individual performance objective. Apart from the flow generated by the noncooperative users, there is also some flow whose routing is controlled by a central network entity, i.e., the manager. For the sake of uniform notation, the manager will also be referred to as user 0. Let $\mathcal{I}_0 = \mathcal{I} \cup \{0\}$.

Let c_l be the capacity of link l, $\mathbf{c} = (c_1, \cdots, c_L)$ the capacity configuration, and $C = \sum_{l \in \mathcal{L}} c_l$ the total capacity of the system of parallel links. We assume that $c_1 \geq \cdots \geq c_L$. Each user $i \in \mathcal{I}_0$ has a throughput demand that is some process with average rate $r^i > 0$. Without loss of generality, we assume that the throughput demands of the noncooperative users satisfy $r^1 \geq r^2 \geq \cdots \geq r^I$. Let $r = \sum_{i \in \mathcal{I}} r^i$ denote the total throughput demand of the noncooperative users, and $R = r + r^0$ the total demand offered to the network. We assume that the system of parallel links can accommodate the total demand, i.e., that $R < C$.

User $i \in \mathcal{I}_0$ ships its flow by splitting its demand r^i over the set of parallel links. Let f_l^i denote the expected flow that user i sends on link l. The user flow configuration $\mathbf{f}^i = (f_1^i, \cdots, f_L^i)$ is called a routing *strategy* of user i and the set $F^i = \{\mathbf{f}^i \in \mathbb{R}^L : 0 \leq f_l^i \leq c_l, l \in \mathcal{L}; \sum_{l \in \mathcal{L}} f_l^i = r^i\}$ of strategies that satisfy the user's demand is called the strategy space of user i. The system flow configuration $\mathbf{f} = (\mathbf{f}^0, \mathbf{f}^1, \cdots, \mathbf{f}^I)$ is called a routing *strategy profile* and takes values in the product strategy space $F = \otimes_{i \in \mathcal{I}_0} F^i$.

The GoS that the flow of user $i \in \mathcal{I}_0$ receives is quantified by means of a cost function $J^i : F \to \mathbb{R}$. $J^i(\mathbf{f})$ is the cost of user i under strategy profile $\mathbf{f}$; the higher $J^i(\mathbf{f})$ is, the lower the GoS provided to the flow of the user. We consider cost functions that are the sum of link cost functions

$$J^i(\mathbf{f}) = \sum_{l \in \mathcal{L}} f_l^i T_l(f_l) \tag{1}$$

where $T_l(f_l)$ is the average delay on link l and depends only on the total flow $f_l = \sum_{i \in \mathcal{I}_0} f_l^i$ on that link. The average delay should be interpreted as a general *congestion cost* per unit of flow, that encapsulates the dependence of the quality of service provided by a finite capacity resource on the total load f_l offered to it (see [12] for a related discussion). In the present paper, we concentrate on congestion costs of the form

$$T_l(f_l) = \begin{cases} (c_l - f_l)^{-1}, & f_l < c_l \\ \infty, & f_l \geq c_l \end{cases} \tag{2}$$

that are typical in various practical routing algorithms [21], [22].[4]

The total cost $J(\mathbf{f})$ of the network depends only on the link flow configuration $(f_1, \cdots, f_L)$

$$J(\mathbf{f}) = \sum_{i \in \mathcal{I}_0} J^i(\mathbf{f}) = \sum_{l \in \mathcal{L}} \frac{f_l}{c_l - f_l}. \tag{3}$$

Since $\sum_l f_l(c_l - f_l)^{-1}$ is a convex function of $(f_1, \cdots, f_L)$, there exists a *unique* link flow configuration $(f_1^*, \cdots, f_L^*)$—with $f_l^* \geq 0$ and $\sum_l f_l^* = R$—that minimizes the total cost. This is the solution of the classical routing optimization problem, where the routing of all flow (generated by both the noncooperative users and the manager) in the network is centrally controlled, and will be referred to as the network optimal link flow configuration, or for simplicity as the *network optimum*. The Kuhn–Tucker optimality conditions [23], imply that $(f_1^*, \cdots, f_L^*)$ is the network optimum if and only if there exists a (Lagrange multiplier) λ^*, such that

[3] More precisely, the threshold decreases as the flow of the users increases.

[4] Note that (2) describes the M/M/1 delay function. Therefore, if we assume that the delay characteristics of each link can be approximated by an M/M/1 queue, $J^i(\mathbf{f})/r^i$ is the average time-delay that the flow of user i experiences under strategy profile $\mathbf{f}$. Similarly if $J(\mathbf{f})$ is the total cost of the network, $J(\mathbf{f})/R$ is the average time-delay experienced by the total flow offered to the network.

for every link $l \in \mathcal{L}$

$$\lambda^* = \frac{c_l}{(c_l - f_l^*)^2}, \quad \text{if} \quad f_l^* > 0 \tag{4}$$

$$\lambda^* \leq \frac{1}{c_l}, \quad \text{if} \quad f_l^* = 0. \tag{5}$$

Let J^* denote the minimal total cost, that is achieved at the network optimum $(f_1^*, \cdots, f_L^*)$. Then, for any strategy profile $\mathbf{f} \in F$, we have $J(\mathbf{f}) \geq J^*$.

A. Noncooperative Users

Each user $i \in \mathcal{I}$ aims to find a routing strategy $\mathbf{f}^i \in F^i$ that minimizes its cost J^i, or equivalently its average time-delay. This optimization problem depends on the routing decisions of the manager and the other users, described by the strategy profile $\mathbf{f}^{-i} = (\mathbf{f}^0, \mathbf{f}^1, \cdots, \mathbf{f}^{i-1}, \mathbf{f}^{i+1}, \cdots, \mathbf{f}^I)$, since J^i is a function of the system flow configuration $\mathbf{f}$.

As already explained, the routing strategy of the manager is fixed, as long as the set of noncooperative users and their throughput demands do not change. Throughout this section we assume that the manager employs strategy $\mathbf{f}^0$, according to some criterion that will be presented in the sequel. Each noncooperative user, on the other hand, adjusts its routing strategy to the actions of the other noncooperative users, in order to minimize its cost. This self-optimizing mode of operation leads to a dynamic behavior that can be modeled as a noncooperative game. Any operating point of the network is a Nash equilibrium of this game, i.e., a strategy profile $\mathbf{f}^{-0}$ of the noncooperative users, from which no user finds it beneficial to unilaterally deviate. These operating points depend on the manager's strategy $\mathbf{f}^0$. Hence, given that the manager employs strategy $\mathbf{f}^0$, strategy profile $\mathbf{f}^{-0} \in F^{-0}$ is a Nash equilibrium of the user routing game if

$$\mathbf{f}^i \in \arg\min_{\mathbf{g}^i \in F^i} J^i(\mathbf{g}^i, \mathbf{f}^{-i}), \quad i \in \mathcal{I}. \tag{6}$$

From the perspective of the users, the manager merely reduces the capacity of each link l by f_l^0. Therefore, the user routing game is equivalent to the routing game in a system of parallel links with capacity configuration $\mathbf{c} - \mathbf{f}^0$. As shown in [8], this routing game has a *unique* Nash equilibrium. Hence, any strategy $\mathbf{f}^0$ of the manager induces a unique Nash equilibrium $\mathbf{f}^{-0}$ of the noncooperative users, that will be denoted by $\mathcal{N}^0(\mathbf{f}^0)$.

Given a strategy profile $\mathbf{f}^{-i}$ of the other users in $\mathcal{I}_0$, the cost of user i, as defined by (1) and (2), is a convex function of its strategy $\mathbf{f}^i$. Hence, the minimization problem in (6) has a unique solution. The Kuhn–Tucker optimality conditions, then, imply that $\mathbf{f}^i$ is the optimal response of user i to $\mathbf{f}^{-i}$ if and only if there exists a (Lagrange multiplier) λ^i, such that, for every link $l \in \mathcal{L}$, we have

$$\lambda^i = \frac{c_l - f_l + f_l^i}{(c_l - f_l)^2}, \quad \text{if} \quad f_l^i > 0 \tag{7}$$

$$\lambda^i \leq \frac{1}{c_l - f_l}, \quad \text{if} \quad f_l^i = 0. \tag{8}$$

Therefore, $\mathbf{f}^{-0} \in F^{-0}$ is the Nash equilibrium of the self-optimizing users induced by strategy $\mathbf{f}^0$ of the manager, if and only if there exist λ^i, $i \in \mathcal{I}$, such that the optimality Conditions (7)–(8) are satisfied for all $i \in \mathcal{I}$.

The function $\mathcal{N}^0 \colon F^0 \to F^{-0}$ that assigns to each strategy of the manager the induced equilibrium of the user routing game is called the *Nash mapping*. From [14, Theorem 3.3], it follows that the Nash mapping is continuous.

B. The Role of the Manager

The manager has knowledge of the noncooperative behavior of the users, that enables it to determine the Nash equilibrium $\mathcal{N}^0(\mathbf{f}^0)$ induced by any routing strategy $\mathbf{f}^0$ that it chooses. Being a central network entity, the manager either has the necessary information available, or can obtain it by monitoring the behavior of the users. This way, the manager can determine a routing strategy of its own flow that gives rise to a Nash equilibrium that is optimal, according to some systemwide efficiency criterion. Therefore, the manager acts as a Stackelberg leader, that imposes its strategy on the self-optimizing users that behave as followers. The presence of sophisticated users that can acquire information about the self-optimizing behavior of the other users and become Stackelberg leaders, in order to optimize their own performance, is in general undesirable. The manager, however, aims at optimizing the overall network performance, thus it plays a social rather than a selfish role.

The goal of the manager is to find a routing strategy of its own flow that drives the system to the network optimum, i.e., a strategy $\mathbf{f}^0$ such that if $\mathbf{f}^{-0} = \mathcal{N}^0(\mathbf{f}^0)$, then $\sum_{i \in \mathcal{I}_0} f_l^i = f_l^*$ for all $l \in \mathcal{L}$. Any such strategy of the manager achieves the minimal total cost J^* and, therefore, leads to the most efficient utilization of network resources. Accordingly, let us introduce the following:

Definition 1: Let $\mathbf{f}^0 \in F^0$ be a strategy of the manager and $\mathbf{f}^{-0} = \mathcal{N}^0(\mathbf{f}^0)$. Strategy $\mathbf{f}^0$ is called *maximally efficient* if it achieves the network optimum, i.e., if $\sum_{i \in \mathcal{I}_0} f_l^i = f_l^*$, $l \in \mathcal{L}$.

An important question is then in place. Maximally efficient strategies of the manager do optimize the overall network performance; but what happens with the performance of the individual users? To address this question, we need to introduce the notion of efficiency at the user level. A standard criterion used in game theory to express this type of efficiency is Pareto efficiency. A strategy profile $\mathbf{f} \in F$ is called *Pareto efficient (or optimal)* if there is no other strategy profile $\hat{\mathbf{f}} \in F$ such that when moving from $\mathbf{f}$ to $\hat{\mathbf{f}}$: 1) all users do at least as well, i.e., $J^i(\hat{\mathbf{f}}) \leq J^i(\mathbf{f})$, for all $i \in \mathcal{I}_0$, and 2) at least one user does strictly better, i.e., there is a user $j \in \mathcal{I}_0$ for which $J^j(\hat{\mathbf{f}}) < J^j(\mathbf{f})$. Clearly, Pareto efficiency is a desirable property for the operating point of the network. Noncooperative equilibria, however, are generically Pareto inefficient [10]. Let us now explain that maximally efficient strategies of the manager drive the network to Pareto efficient operating points.

Assume that $\mathbf{f} = (\mathbf{f}^0, \mathcal{N}^0(\mathbf{f}^0))$ is the operating point induced by a maximally efficient strategy $\mathbf{f}^0$ of the manager. To see that $\mathbf{f}$ is Pareto efficient, assume that there exists another strategy profile $\hat{\mathbf{f}} \neq \mathbf{f}$ that satisfies Conditions 1) and 2) above. Then $J(\hat{\mathbf{f}}) = \sum_{i \in \mathcal{I}_0} J^i(\hat{\mathbf{f}}) < \sum_{i \in \mathcal{I}_0} J^i(\mathbf{f}) = J(\mathbf{f}) = J^*$, which is a contradiction, since J^* is the cost achieved at the network optimum. Therefore we have the following:

Proposition 1: A maximally efficient strategy $\mathbf{f}^0$ of the manager induces a Pareto efficient operating point $(\mathbf{f}^0, \mathcal{N}^0(\mathbf{f}^0))$.

Continuity of the Nash mapping implies that $J(\mathbf{f}^0, \mathcal{N}^0(\mathbf{f}^0))$ is continuous in $\mathbf{f}^0 \in F^0$, thus it attains its minimum in the compact set F^0. Therefore, an *optimal* strategy of the manager always exists. Existence of a maximally efficient strategy, however, cannot be guaranteed in general. Evidently, if a maximally efficient strategy exists, then it is an optimal strategy of the manager. In the following sections, we derive necessary and sufficient conditions that guarantee existence of a maximally efficient strategy of the manager. Before we proceed with the analysis, let us present an informal summary of the main results.

III. Outline of Results

1) In the special case of a single user, the manager can always enforce the network optimum, and its maximally efficient strategy is specified explicitly.
2) In the general case of any number of users, the manager can enforce the network optimum if and only if its demand is higher than some threshold $\underline{r}^0$, in which case the manager's maximally efficient strategy is specified explicitly.
3) The threshold $\underline{r}^0$ is feasible, in the sense that the total demand of the users plus $\underline{r}^0$ is lower than the total capacity of the network. Thus, for every set of users (whose total demand r is less than the total capacity C) there are managers that can enforce the network optimum.
4) In heavily loaded networks it is "easy" for the manager to enforce the network optimum (i.e., the threshold $\underline{r}^0$ is small).
5) As the number of users increases, it becomes harder for the manager to enforce the network optimum (i.e., the threshold $\underline{r}^0$ increases).
6) The higher the difference in the throughput demands of any two users, the easier it becomes for the manager to enforce the network optimum.

IV. Preliminary Structural Results

The structure of the Nash equilibrium in a system of parallel links shared by I noncooperative users has been investigated in [8] and [14]. The results of these references can be readily applied to characterize the structure of the network optimum and the Nash mapping. In this section we briefly present the related results without proofs.

Let us first consider the network optimum $(f_1^*, \cdots, f_L^*)$. The flow f_l^* on link l, is decreasing in the link number $l \in \mathcal{L}$. Therefore, there exists some link L^*, such that $f_l^* > 0$ for $l \leq L^*$ and $f_l^* = 0$ for $l > L^*$. The threshold L^* is determined by

$$G_{L^*} < R \leq G_{L^*+1} \tag{9}$$

where

$$G_l = \sum_{n=1}^{l-1} c_n - \sqrt{c_l} \sum_{n=1}^{l-1} \sqrt{c_n}, \quad l = 2, \cdots, L \tag{10}$$

and $G_1 = 0$, $G_{L+1} = \sum_{n=1}^{L} c_n = C$. Note that $c_l \geq c_{l+1}$ implies that $G_l \leq G_{l+1}$ for all $l \in \mathcal{L}$.

Using the optimality Conditions (4)–(5), it can be easily verified that

$$c_l - f_l^* \geq c_{l+1} - f_{l+1}^*, \quad l = 1, \cdots, L-1 \tag{11}$$

with equality holding if and only if $c_l = c_{l+1}$. Moreover, writing (4) as $\sqrt{\lambda^*}(c_l - f_l^*) = \sqrt{c_l}$, and summing over any set of links $A \subseteq \{1, \cdots, L^*\}$, we have

$$\lambda^* = \left[\frac{\sum_{l \in A} \sqrt{c_l}}{\sum_{l \in A}(c_l - f_l^*)}\right]^2, \quad A \subseteq \{1, \cdots, L^*\}. \tag{12}$$

Finally, the network optimum $(f_1^*, \cdots, f_L^*)$ is given by [14]

$$f_l^* = \begin{cases} c_l - \left(\sum_{n=1}^{L^*} c_n - R\right) \frac{\sqrt{c_l}}{\sum_{n=1}^{L^*} \sqrt{c_n}}, & l \leq L^* \\ 0, & l > L^*. \end{cases} \tag{13}$$

Let us now consider the Nash equilibrium $\mathbf{f}^{-0} = \mathcal{N}^0(\mathbf{f}^0)$ of the users that is induced by strategy $\mathbf{f}^0$ of the manager. In order to characterize the structure of $\mathbf{f}^{-0}$, it suffices to determine the best reply $\mathbf{f}^i$ of user $i \in \mathcal{I}$ to the strategies of the other users and the manager that are described by $\mathbf{f}^{-i}$. For any link l, let $c_l^i = c_l - \sum_{j \in \mathcal{I}_0 \setminus \{i\}} f_l^j$ denote the residual capacity of the link as seen by user i. Then, $\mathbf{f}^i$ can be determined as the network optimum for a system of parallel links with capacity configuration $(c_1^i, \cdots, c_L^i)$. Therefore, *assuming* that

$$c_l^i \geq c_{l+1}^i, \quad l = 1, \cdots, L-1 \tag{14}$$

the flow f_l^i is decreasing in the link number $l \in \mathcal{L}$. This implies that there exists some link L^i, such that $f_l^i > 0$ for $l \leq L^i$ and $f_l^i = 0$ for $l > L^i$. The threshold L^i is determined by

$$G_{L^i}^i < r^i \leq G_{L^i+1}^i \tag{15}$$

where, similarly to (10)

$$G_l^i = \sum_{n=1}^{l-1} c_n^i - \sqrt{c_l^i} \sum_{n=1}^{l-1} \sqrt{c_n^i}, \ l = 2, \cdots, L \tag{16}$$

and $G_1^i = 0$, $G_{L+1}^i = \sum_{n=1}^{L} c_n^i = C - (R - r^i)$. (14) implies that $G_l^i \leq G_{l+1}^i$ for all $l \in \mathcal{L}$.

Similarly to (13), the best reply $\mathbf{f}^i$ of user i to strategy profile $\mathbf{f}^{-i}$ of the other users in $\mathcal{I}_0$ is given by

$$f_l^i = \begin{cases} c_l^i - \left(\sum_{m=1}^{L^i} c_m^i - r^i\right) \frac{\sqrt{c_l^i}}{\sum_{m=1}^{L^i} \sqrt{c_m^i}}, & l \leq L^i. \\ 0, & l > L^i \end{cases} \tag{17}$$

Equations (16) and (17) indicate that the information user i needs to determine its best reply $\mathbf{f}^i$ to any strategy profile $\mathbf{f}^{-i}$ is the residual capacity c_l^i seen by the user on every link $l \in \mathcal{L}$, and not a detailed description of $\mathbf{f}^{-i}$. In practice, information about the residual capacities can be acquired by measuring the link delays through an appropriate estimation technique.

V. Single-Follower Stackelberg Routing Game

In this section we consider the simplest case of a Stackelberg routing game, where the network is shared by a single self-optimizing user $(I = 1)$ and the manager. The simplicity of this model will allow us to elucidate both the intuition behind the structure of the manager's maximally efficient strategy and the methodology to derive it. Moreover, the results of this section provide the foundation for the analysis of the general Stackelberg routing game that will be carried out in the following section. The proofs of all results in this section are presented in Appendix A.

First we investigate the structure of a maximally efficient strategy $\mathbf{f}^0$ of the manager, provided that one exists. Then, we establish existence of a maximally efficient strategy. Before we proceed, let us define

$$H_l = \sum_{n=1}^{l-1} f_n^* - \frac{f_l^*}{c_l} \sum_{n=1}^{l-1} c_n, \quad l = 2, \cdots, L \tag{18}$$

and $H_1 = 0$, $H_{L+1} = \sum_{n=1}^{L} f_n^* = R$. Using (4) and (10), it is easy to see that

$$H_l = \begin{cases} G_l/\sqrt{\lambda^* c_l}, & l \leq L^* \\ R, & l > L^* \end{cases} \tag{19}$$

thus

$$H_l \leq H_{l+1}, \quad l = 1, \cdots, L. \tag{20}$$

The following lemma shows that if a maximally efficient strategy of the manager exists, then it is unique.

Lemma 1: In the single-follower Stackelberg routing game, if there exists a maximally efficient strategy of the manager, then it is unique and is given by

$$f_l^0 = \begin{cases} c_l \frac{\sum_{n=1}^{L^1} f_n^* - r^1}{\sum_{n=1}^{L^1} c_n}, & l \leq L^1 \\ f_l^*, & l > L^1 \end{cases} \tag{21}$$

where L^1 is determined by

$$H_{L^1} < r^1 \leq H_{L^1+1}. \tag{22}$$

If a maximally efficient strategy $\mathbf{f}^0$ of the manager exists, then, (21) implies that the best reply $\mathbf{f}^1$ of the follower is

$$f_l^1 = f_l^* - f_l^0 = \begin{cases} f_l^* - c_l \frac{\sum_{n=1}^{L^1} f_n^* - r^1}{\sum_{n=1}^{L^1} c_n}, & l \leq L^1 \\ 0, & l > L^1. \end{cases} \tag{23}$$

Therefore, $\{1, \cdots, L^1\}$ is the set of links over which the follower sends its flow, when the manager implements $\mathbf{f}^0$. According to (21), the manager: 1) sends flow f_l^*, on every link l that will not receive any flow from the follower and 2) splits the rest of its flow among the links that will receive flow from the follower, proportionally to their capacities.

To establish existence of the maximally efficient strategy of the manager, it suffices to show that $\mathbf{f}^0$ given by (21) and (22) is such that:

1) $\mathbf{f}^0$ is an admissible strategy of the manager, i.e., $f_l^0 \geq 0$, $l \in \mathcal{L}$, and $\sum_{l \in \mathcal{L}} f_l^0 = r^0$;
2) $\mathbf{f}^1$, with f_l^1 given by (23) for $l \leq L^1$, and $f_l^1 = 0$ for $l > L^1$, is the best reply of the follower to $\mathbf{f}^0$, i.e., $\mathbf{f}^1 = \mathcal{N}^0(\mathbf{f}^0)$.

This is established in the following theorem that gives the main result of this section.

Theorem 1: In the single-follower Stackelberg routing game, there exists a unique maximally efficient strategy $\mathbf{f}^0$ of the leader that is described by (21) and (22).

The above theorem indicates that, for a single follower, the leader can always enforce the network optimum, independently of the relative sizes in terms of throughput demands of the leader and the follower. As will be seen in the following section, this might not be the case in the presence of multiple self-optimizing users.

VI. Multifollower Stackelberg Routing Game

Let us now proceed with the general Stackelberg routing game, where an arbitrary number I of self-optimizing users share the system of parallel links. The following lemma describes the maximally efficient strategy of the manager—provided that one exists—as well as the corresponding Nash equilibrium of the noncooperative users. Later, we will derive necessary and sufficient conditions that guarantee existence of a maximally efficient strategy of the manager. The proofs of all results in this section are presented in Appendix B.

Lemma 2: In a multifollower Stackelberg routing game, if there exists a maximally efficient strategy $\mathbf{f}^0$ of the leader, then it is unique and is given by

$$f_l^0 = c_l \sum_{i \in \mathcal{I}_l} \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} - (I_l - 1) f_l^*, \quad l \in \mathcal{L} \tag{24}$$

where, for every user $i \in \mathcal{I}$, L^i is determined by

$$H_{L^i} < r^i \leq H_{L^i+1} \tag{25}$$

and for every link $l \in \mathcal{L}, \mathcal{I}_l = \{i \in \mathcal{I} : l \leq L^i\}$ and $I_l = |\mathcal{I}_l|$. In that case, the equilibrium strategy $\mathbf{f}^i$ of user $i \in \mathcal{I}$ is described by

$$f_l^i = \begin{cases} f_l^* - c_l \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n}, & l \leq L^i \\ 0, & l > L^i. \end{cases} \tag{26}$$

Conversely, if $\mathbf{f}^0$ described by (24) and (25) is an admissible strategy of the leader, then it is its maximally efficient strategy.

Note that, if a maximally efficient strategy of the manager exists, then the induced Nash equilibrium of the followers has precisely the same structure with the best reply of the follower in the single-follower case, that is given by (23) and (22).

Remarks:

1) $\mathbf{f}^0$ given by (24) and (25) might fail to be an admissible strategy of the leader; it merely decreases/increases the capacity of link $l \in \mathcal{L}$ when f_l^0 is positive/negative. From the proof of the lemma in Appendix B, it follows that, even if $\mathbf{f}^0$ is nonadmissible, $\mathbf{f}^{-0}$ with $\mathbf{f}^i$ given by (26) for $i \in \mathcal{I}$ is the induced Nash equilibrium of the followers.
2) Under (26), $\{1, \cdots, \mathcal{L}^i\}$ is the set of links that receive flow from follower $i \in \mathcal{I}$. Thus, $\mathcal{I}_l$ is precisely theset of followers that send flow on link $l \in \mathcal{L}$. Since

$H_1 = 0 < r^i$, $i \in \mathcal{I}$, all users send flow on link 1, that is, $\mathcal{I}_1 = \mathcal{I}$.

3) Since $r^i \geq r^{i+1}$, (25) implies $L^i \geq L^{i+1}$ for all $i < I$, and $\mathcal{I}_{l+1} \subseteq \mathcal{I}_l$ for all $l < L$. Furthermore, since $r^i \leq R = H_{L^*+1}$, (25) implies that $L^i \leq L^*$, $i \in \mathcal{I}$.

Let us now investigate the admissibility of $\mathbf{f}^0$. To this end, observe that

$$\begin{aligned}\sum_{l=1}^{L} f_l^0 &= \sum_{l=1}^{L}\sum_{i\in\mathcal{I}_l} c_l \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} - \sum_{l=1}^{L} I_l f_l^* + \sum_{l=1}^{L} f_l^* \\ &= \sum_{i=1}^{I}\sum_{l=1}^{L^i} c_l \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} - \sum_{l=1}^{L} I_l f_l^* + r^0 + \sum_{i=1}^{I} r^i \\ &= \sum_{i=1}^{I}\sum_{n=1}^{L^i} f_n^* - \sum_{l=1}^{L} I_l f_l^* + r^0 = r^0 \end{aligned} \tag{27}$$

since $\sum_{i=1}^{I}\sum_{n=1}^{L^i} f_n^* = \sum_{l=1}^{L} I_l f_l^*$. Thus, $\mathbf{f}^0$ is admissible if and only if $f_l^0 \geq 0$, for all $l \in \mathcal{L}$. The following lemma implies that this condition can be relaxed to $f_1^0 \geq 0$.

Lemma 3: Consider the (possibly nonadmissible) strategy $\mathbf{f}^0$ of the leader, that is given by (24) and (25). For every link $l > 1$, we have

$$f_l^0 < 0 \Rightarrow f_{l-1}^0 < 0.$$

The previous lemma, together with Lemma 2, implies that a maximally efficient strategy of the leader exists if and only if f_1^0 given by (24) is nonnegative. The following lemma shows that f_1^0 is an increasing function of the throughput demand r^0 of the leader. This monotonicity property is used in the sequel to establish that a maximally efficient strategy of the leader exists if and only if its demand is sufficiently large.

Lemma 4: Let f_1^0 be as in (24). Then, f_1^0 is a continuous increasing function of the throughput demand $r^0 \in [0, C - r]$ of the leader.

Remark: If $r^0 = C - r$, then $R = C$ and the network becomes saturated. Allowing, however, r^0 to take this value is a mere technicality that will be used in the proof of the following theorem. Note that when the network is saturated, $f_l^* = c_l$ for every link $l \in \mathcal{L}$.

We are now ready to prove the main result of this section that is given in the following.

Theorem 2: There exists some $\underline{r}^0$, with $0 \leq \underline{r}^0 < C - r$, such that the leader in a multifollower Stackelberg routing game can enforce the network optimum if and only if its throughput demand r^0 satisfies $\underline{r}^0 \leq r^0 < C - r$. Then, the maximally efficient strategy of the leader is given by (24) and (25).

As seen by the proof of the theorem in Appendix B, the threshold $\underline{r}^0$ of the leader is the unique solution of the equation "$f_1^0(r^0) = 0$" in $r^0 \in [0, C - r)$. Since f_1^0 is an increasing function of r^0, this equation can be easily solved using standard numerical techniques.

The above theorem implies that, for any finite set of followers with total demand r that does not exceed the total capacity C of the system, there is always a (feasible) leader, with $\underline{r}^0 \leq r^0 < C - r$, that can enforce the network optimum. Moreover, when $r \to C$, we have $\underline{r}^0 \to 0$, meaning that in heavily loaded networks it suffices to control just a small portion of the flow in order to drive the system into the network optimum.

The threshold $\underline{r}^0$ on the leader's throughput demand depends on the number and the throughput demands of the followers. This dependence is investigated in the following section.

A. Properties of the Leader Threshold $\underline{r}^0$

Let us first examine the dependence of $\underline{r}^0$ on the number of followers, when their throughput demand r is fixed. To simplify the formulation of the problem, we concentrate on followers with identical throughput demands, i.e., with $r^i = r^j$ for all $i, j \in \mathcal{I}$. This class of followers will be referred to as *identical followers*, and the special structure of their Nash equilibrium has been investigated in [8]. The following proposition shows that as the number of followers increases, it becomes harder for the leader to enforce the network optimum.

Proposition 2: Suppose that the followers are identical and their total throughput demand r is fixed. Then, the leader threshold $\underline{r}^0$ is nondecreasing in the number of followers.

Let us now concentrate on the dependence of $\underline{r}^0$ on differences of the demands of the followers, when their total throughput demand r is fixed. The following proposition shows that the higher the difference in the throughput demand of any two followers, the easier it becomes for the leader to enforce the network optimum.

Proposition 3: Suppose that the total throughput demand r of the followers is fixed. Then, for any two followers j and k, the leader threshold $\underline{r}^0$ is nonincreasing in $|r^j - r^k|$. Therefore, $\underline{r}^0$ attains its maximum value when all followers are identical.

Let us now demonstrate the properties of $\underline{r}^0$, established in the previous propositions, by means of a numerical example. We consider a system of parallel links with capacity configuration $\mathbf{c} = (12, 7, 5, 3, 2, 1)$, shared by I identical followers with total demand r. The threshold $\underline{r}^0$ of the leader is depicted in Fig. 1 as a function of r, for various values of I. In the same figure, we also show the saturation line "$\underline{r}^0 + r = C$." From the figure, one can see that $\underline{r}^0$ always lies below the saturation line, in accordance with Theorem 2. Furthermore, $\underline{r}^0$ increases with the number of users.

From the same figure, we observe that in the light load region (i.e., when the total demand r of the followers is low compared to the total capacity C) $\underline{r}^0$ increases with r, that is, the higher the demand of the followers, the more difficult it becomes for the leader to drive the system to the network optimum. In the moderate and heavy load regions, on the other hand, $\underline{r}^0$ is decreasing in r. This behavior has been explained in the discussion following Theorem 2.

VII. Practical Considerations

A. Efficiency

Let us now demonstrate the efficiency of the proposed management scheme, by means of a numerical example. Consider a system of parallel links with capacity configuration $\mathbf{c} = (12, 7, 5, 3, 2, 1)$, shared by $I = 100$ identical self-

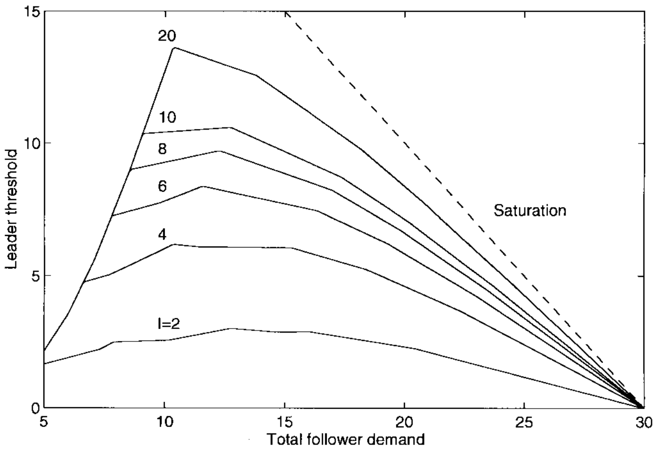


Fig. 1. Leader threshold as a function of total follower demand.

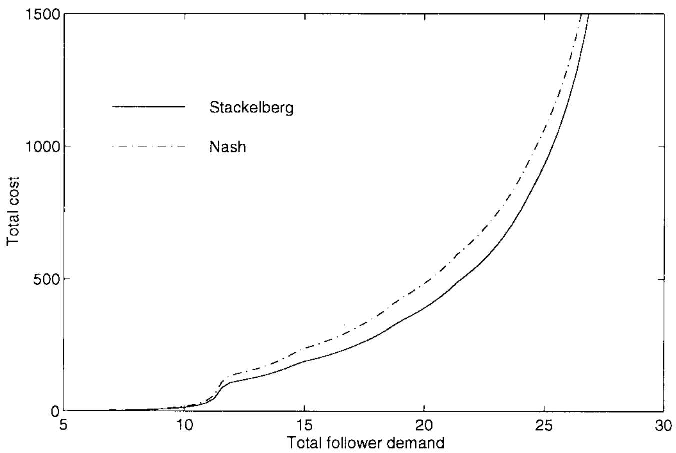


Fig. 2. Total cost as a function of total follower demand.

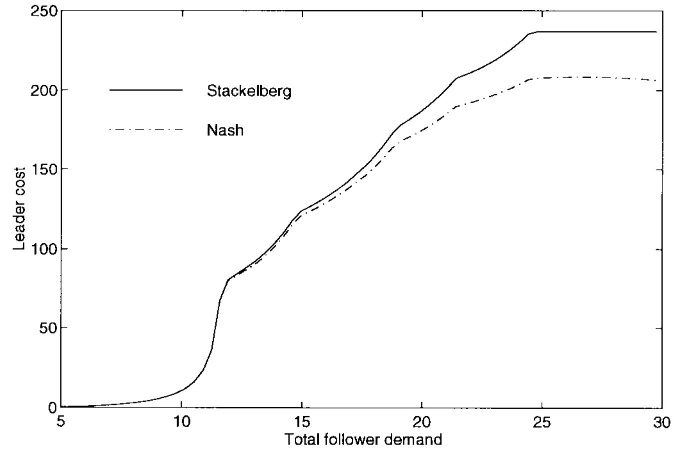


Fig. 3. Leader cost as a function of total follower demand.

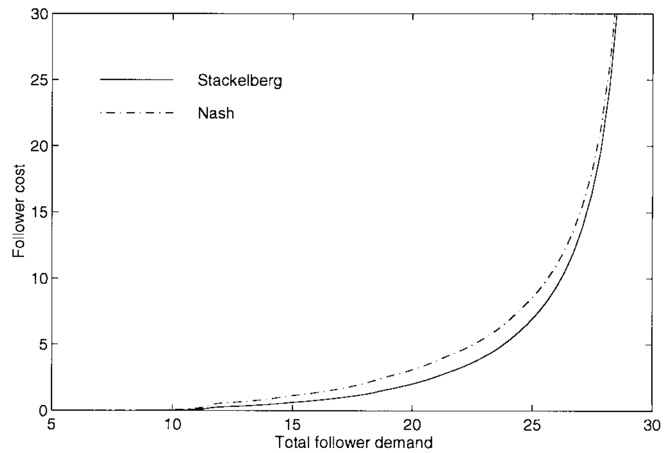


Fig. 4. Follower cost as a function of total follower demand.

optimizing users, with total demand r, and the manager. We investigate the performance of the network under two scenaria: the "Stackelberg" scenario and the "Nash" scenario. In both scenaria, given the total follower demand r, we compute the threshold $\underline{r}^0$ that enables the manager to enforce the network optimum and we assume that its throughput demand is equal to that threshold, that is, $r^0 = \underline{r}^0$. Under the Stackelberg scenario, the manager implements its maximally efficient strategy and drives the system to the network optimum, while under the Nash scenario, it behaves as another noncooperative user. We compare the total cost (Fig. 2), the cost of the manager (Fig. 3) and the follower cost (Fig. 4) under the two scenaria, for different values of the total follower demand. Fig. 5 shows the change (percent) in the total, leader and follower costs, when we move from the Nash to the Stackelberg scenario.

Fig. 2 shows that the network performance is always better under the Stackelberg scenario. More specifically, the improvement in the total cost is more that 20%, as long as the follower demand does not exceed 2/3 of the total capacity of the network. Note that the total cost in the Stackelberg scenario corresponds to the network optimum. Therefore, Fig. 2 demonstrates the inefficiency of noncooperative equilibria for the routing model under consideration.

Fig. 3 indicates that the cost of the manager (the average delay experienced by its flow) is always higher under the Stackelberg scenario, while Fig. 4 shows that the cost of each follower is lower under the same scenario. The increase of the manager cost, however, does not exceed 7% as long as the follower demand does not exceed 2/3 of the total network capacity, while the improvement in the total cost is more than 20% and in the follower cost more than 36%.

B. Scalability

Lemma 2 indicates that in order to determine its maximally efficient strategy, the manager must have information about the throughput demand r^i of every user $i \in \mathcal{I}$, and about the network optimum $(f_1^*, \cdots, f_L^*)$. The network optimum can be readily computed from (13) and (9), given the total load R offered to the network. Hence, the manager needs only information about the throughput demand of every user. In many networks of interest, user flows are accepted by means of some admission control mechanism. This involves a negotiation phase, where the user has to declare certain traffic parameters for its flow, one of which is typically the average rate r^i. Therefore, the manager has readily available the information it needs in order to implement its maximally efficient strategy.[5] Each time a user arrives to or departs from the network, the manager can simply adjust its strategy to

[5] Alternatively, the manager can obtain estimates of the average rates by monitoring the behavior of the users.

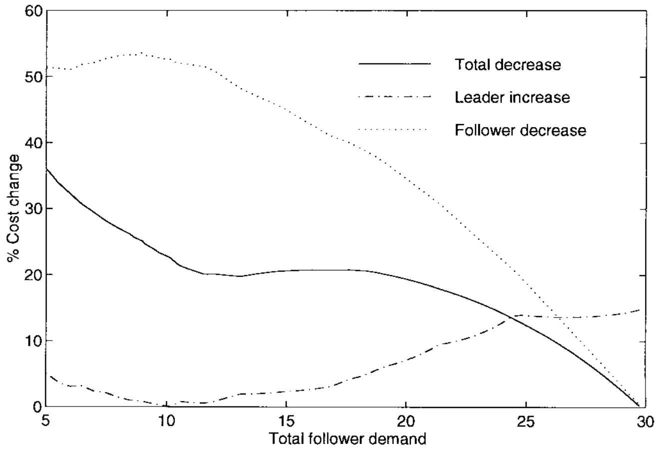


Fig. 5. Change in the total/leader/follower costs versus r.

the maximally efficient one, using the information about the throughput demand of that user. In that sense, the proposed mechanism of enforcing the network optimum by means of the manager's routing strategy is scalable.

The proposed methodology is applicable for many intranets (e.g., institutional or corporate networks), where devising efficient management schemes in order to provide good GoS to the users is of central importance. One might argue, though, that the methodology does not apply to very-large-scale networks, such as internets, where not only it is difficult to keep track of a very large number of users, but also there are no provisions for guaranteeing good quality of service. Two comments are then in place. First, what we refer to as a "user" is not necessarily an individual end user of the network. For metropolitan area networks, for example, users can be institutions or corporations and, therefore, their number is "manageable." Second, internetworking is currently moving toward a commercial Internet that is available to the general public through private service providers [24]. Large Internet service providers (such as long distance carriers) will probably build their own internets and try to attract customers by guaranteeing high quality of service within the boundaries of their internets (while, at the same time, providing gateways to the public Internet). Efficient management schemes will play a central role in achieving this goal.

C. Achieving the Threshold $\underline{r}^0$

An important question that arises from the present work is whether and how the manager can satisfy the necessary and sufficient condition that allows it to drive the system to the network optimum. As indicated by Proposition 2, the minimum demand $\underline{r}^0$ that enables the manager to enforce the network optimum increases with the number of noncooperative users. Therefore, one way to achieve this threshold is to provide incentives to "small" users to join "larger" (but still self-optimizing) network entities, such as virtual networks (VN's). It is worth noting that, while bifurcated routing might seem impractical in the single (small) user case, a VN control entity can implement (optimal) bifurcation by routing the flow of different VN users over its various paths.

An alternative way to achieve the loading threshold $\underline{r}^0$ is to provide incentives to the (noncooperative) users to join a "social" entity (e.g., a "social" VN), that is, one whose flow is directly controlled by the manager. This way, not only the number of noncooperative users is reduced, but also the total flow controlled by the manager is increased.

A key question is, then, what are the possible incentives that would persuade a user to join such larger network entities. One way to achieve this is through appropriate pricing mechanisms. A user may decide to join a VN controlled by the manager, for example, provided that lower prices would compensate for losing control of its flow that might result in a degradation in the GoS that it receives (as indicated by the example in the previous subsection). Moreover, the manager has the flexibility to provide different GoS to the various VN's (or users) it controls, by routing their flow over different paths, while still implementing the maximally efficient strategy for the total flow it controls. The manager can, then, charge a VN (or user) according to the GoS that it receives. Since pricing is one of the key factors for the deployment of future broadband/multimedia networks, investigating such mechanisms is a challenging problem for future research.

VIII. Conclusion

The practical inability to achieve global cooperation in many modern networking environments, typically results in inefficient use of network resources. This situation might be prohibitive for future broadband networks that are expected to support numerous resource consuming applications, such as multimedia. In recent years, a number of methods have been proposed to overcome this problem. These methods improve the network performance either through proper design of the resource configuration and/or the service disciplines of the network, or by introducing some "external" component such as prices.

We proposed a new method for improving the performance of noncooperative networks. This approach calls for the intervention of a social agent, namely the network manager, that tries to optimize the network performance, through the limited control that it routinely employs during the run time phase of the network. Specifically, we considered a network manager that acts as a Stackelberg leader. The manager controls only part of the network flow, and is cognizant of the presence of noncooperative users. Considering a system of parallel links, we showed that, by controlling just a small portion of the network flow, the operating point of the system can often be driven into the network optimum. For situations that the manager does not control enough flow to enforce the network optimum, our analysis provides guidelines for actions that it can take in order to meet the required threshold (e.g., providing monetary incentives to users to join virtual networks that are controlled by the manager).

In practical terms, an important advantage of the proposed management scheme is that it can be readily implemented through appropriate routing of some centrally controlled network flow. This is to be contrasted with other recent proposals that require changes in technology and/or policy making in modern networks. Deployment of recent proposals for sophisticated pricing schemes, for instance, will require changes both in the accounting infrastructure and the policy making process in current networks.

It should be noted that our analysis depends on the specific structure of the model. The extent to which these results can be generalized is an important subject for further research. Nonetheless, the ability to obtain efficient strategies for simple networking models inherently has important implications. We indicated, for example, that systems of parallel links appropriately model scenaria that become common in modern networking. Indeed, current practices tend to decrease the degrees of freedom in networks, as is the case, for example, when bandwidth is separated among virtual paths. The present work indicates that such practices make the network less vulnerable to the deficiencies of noncooperation. This is yet a further indication of the potential benefit of decoupling complex structures in a network.

Appendix A
Proofs of Results in Section V

Proof of Lemma 1: Suppose that there exists a maximally efficient strategy $\mathbf{f}^0$ of the manager and let $\mathbf{f}^1 = \mathcal{N}^0(\mathbf{f}^0)$ be the best reply of the follower. Then

$$f_l = f_l^0 + f_l^1 = f_l^*, \quad l \in \mathcal{L}. \tag{28}$$

Let us first show that the flow f_l^1 the follower sends on link l is decreasing in the link number $l \in \mathcal{L}$. Assume by contradiction that, for some n, we have $0 \le f_n^1 < f_{n+1}^1$. Then, the optimality Conditions (7)–(8) imply that

$$\frac{1}{c_{n+1} - f_{n+1}^*} + \frac{f_{n+1}^1}{(c_{n+1} - f_{n+1}^*)^2} \le \frac{1}{c_n - f_n^*} + \frac{f_n^1}{(c_n - f_n^*)^2}$$

which is a contradiction, since $c_n - f_n^* \ge c_{n+1} - f_{n+1}^*$ [by (11)] and $f_{n+1}^1 > f_n^1$ (by assumption). Therefore, there exists some link L^1, such that $f_l^1 > 0$ for $l \le L^1$ and $f_l^1 = 0$ for $l > L^1$, that is, the follower sends its flow precisely over the links in $\{1, \cdots, L^1\}$. Furthermore, (11) and $f_l^1 \ge f_{l+1}^1$ imply that for any link l, we have $c_l^1 = c_l - f_l^* + f_l^1 \ge c_{l+1} - f_{l+1}^* + f_{l+1}^1 = c_{l+1}^1$, that is, the residual link capacities as seen by the follower preserve the order of the link capacities themselves. Hence, the threshold L^1 is determined by (15), with $i = 1$, as explained in Section IV. In view of (28), it is evident that $L^1 \le L^*$.

The optimality Conditions (4)–(5) for $(f_1^*, \cdots, f_L^*)$ and (7)–(8) for $\mathbf{f}^1$ imply that $\frac{c_l^1}{c_m^1} = \frac{c_l}{c_m} = [\frac{c_l - f_l^*}{c_m - f_m^*}]^2$, for all $l, m \in \{1, \cdots, L^1\}$. Taking $m = 1$, we have $f_l^1 = f_l^* - \frac{c_l}{c_1}(f_1^* - f_1^1)$, $l = 1, \cdots, L^1$, which, together with $\sum_{l=1}^{L^1} f_l^1 = r^1$, gives

$$f_l^1 = f_l^* - \frac{c_l}{\sum_{n=1}^{L^1} c_n} \left[\sum_{n=1}^{L^1} f_n^* - r^1 \right], \quad l = 1, \cdots, L^1. \tag{29}$$

Then, (21) is immediate from (28) and (29).

We now proceed with the proof of (22). Since the follower sends its flow precisely over the links in $\{1, \cdots, L^1\}$, we have $G_{L^1}^1 < r^1 \le G_{L^1+1}^1$. The residual capacity seen by the follower on any link $l \le L^1$ is

$$c_l^1 = c_l \frac{\sum_{n=1}^{L^1}(c_n - f_n^*) + r^1}{\sum_{n=1}^{L^1} c_n}, \quad l = 1, \cdots, L^1 \tag{30}$$

according to (21).

Let us first show that $H_{L^1} < r^1$. Using (16) and (30), $r^1 > G_{L^1}^1$ is equivalent to

$$r^1 > \left[\sum_{n=1}^{L^1-1} c_n - \sqrt{c_{L^1}} \sum_{n=1}^{L^1-1} \sqrt{c_n} \right] \frac{\sum_{n=1}^{L^1}(c_n - f_n^*) + r^1}{\sum_{n=1}^{L^1} c_n}$$

or

$$r^1 \sqrt{c_{L^1}} \sum_{n=1}^{L^1} \sqrt{c_n} > \left[\sum_{n=1}^{L^1-1} c_n - \sqrt{c_{L^1}} \sum_{n=1}^{L^1-1} \sqrt{c_n} \right] \sum_{n=1}^{L^1} (c_n - f_n^*). \tag{31}$$

Since $L^1 \le L^*$, taking $A = \{1, \cdots, L^1\}$ in (12), we get

$$\sqrt{\lambda^*} = \frac{\sum_{n=1}^{L^1} \sqrt{c_n}}{\sum_{n=1}^{L^1}(c_n - f_n^*)}. \tag{32}$$

Thus, (31) is equivalent to

$$r^1 > \frac{\sum_{n=1}^{L^1-1} c_n - \sqrt{c_{L^1}} \sum_{n=1}^{L^1-1} \sqrt{c_n}}{\sqrt{\lambda^* c_{L^1}}} = \frac{G_{L^1}}{\sqrt{\lambda^* c_{L^1}}} = H_{L^1}.$$

Let us now proceed to show that $r^1 \le H_{L^1+1}$. If $f_{L^1+1}^* = 0$, then $L^* = L^1$ and $H_{L^1+1} = R > r^1$, by (19). Therefore, we concentrate on the case where $f_{L^1+1}^* > 0$. Using (16) and (30), and after some algebraic manipulation, $r^1 \le G_{L^1+1}^1$ is equivalent to

$$r^1 \le \frac{1}{c_{L^1+1} - f_{L^1+1}^*} \times \left[\frac{\sum_{n=1}^{L^1}(c_n - f_n^*)}{\sum_{n=1}^{L^1} \sqrt{c_n}} \right]^2 \sum_{n=1}^{L^1} c_n - \sum_{n=1}^{L^1} (c_n - f_n^*)$$

and, using (32), equivalent to

$$r^1 \le \frac{1}{c_{L^1+1} - f_{L^1+1}^*} \frac{1}{\lambda^*} \sum_{n=1}^{L^1} c_n - \frac{\sum_{n=1}^{L^1} \sqrt{c_n}}{\sqrt{\lambda^*}}. \tag{33}$$

Since $f_{L^1+1}^* > 0$, (7) gives $c_{L^1+1} - f_{L^1+1}^* = \sqrt{c_{L^1+1}/\lambda^*}$, and (33) is equivalent to

$$r^1 \le \frac{\sum_{n=1}^{L^1} c_n - \sqrt{c_{L^1+1}} \sum_{n=1}^{L^1} \sqrt{c_n}}{\sqrt{\lambda^* c_{L^1+1}}} = H_{L^1+1} \tag{34}$$

and this concludes the proof of (22).

Since H_l is independent from the manager's strategy $\mathbf{f}^0$, for all l, condition (22) is also independent of $\mathbf{f}^0$. Furthermore, in view of (20), it determines the threshold L^1 uniquely. Therefore, if a maximally efficient strategy of the manager exists, then it is unique and is given by (21) and (22). □

Proof of Theorem 1: From $H_{L^1} < r^1 \leq R = H_{L^*+1}$, we conclude that $L^1 \leq L^*$. From $H_{L^1+1} \geq r^1$ and (18) we have $\sum_{n=1}^{L^1} f_n^* \geq r^1$. Therefore, $f_l^0 \geq 0$ for $l \leq L^1$, from (21). Furthermore, it is easy to verify that $\sum_{l \in \mathcal{L}} f_l^0 = \sum_{l \in \mathcal{L}} f_l^* - r^1 = r^0$. Thus, $\mathbf{f}^0$ is an admissible strategy of the manager.

We now proceed to show that $\mathbf{f}^1$ given by (23) is the best reply of the follower to $\mathbf{f}^0$. For all $l \leq L^1$, (20) and (22) imply $(c_l / \sum_{n=1}^{l} c_n)(\sum_{n=1}^{l} f_n^* - r^1) < f_l^*$, which together with (21) gives $f_l^0 < f_l^*$. Thus, $f_l^1 = f_l^* - f_l^0 > 0$. Moreover, $\sum_{l \in \mathcal{L}} f_l^1 = r^1$, by (23). Hence, $\mathbf{f}^1 \in F^1$. Let us now show that the residual capacities seen by the follower preserve the original link ordering, that is, they satisfy (14). For $l > L^1$, this is immediate from (11). The same is true for $l \leq L^1$, in view of (30). Finally, for $l = L^1$, we have $c_{L^1}^1 = c_{L^1} - f_{L^1}^0 \geq c_{L^1} - f_{L^1}^* \geq c_{L^1+1} - f_{L^1+1}^* = c_{L^1+1} - f_{L^1+1}^0 = c_{L^1+1}^1$, where the first inequality follows from $f_{L^1}^0 \leq f_{L^1}^*$, and the second from (11). Thus, inequality (14) holds. This implies that the best reply of the follower to $\mathbf{f}^0$ has the threshold structure of $\mathbf{f}^1$, where the respective threshold, say N^1, is determined by $G_{N^1}^1 < r^1 \leq G_{N^1+1}^1$. To show $N^1 = L^1$, it suffices to show that $G_{L^1}^1 < r^1 \leq G_{L^1+1}^1$.

Recall that $H_{L^1} < r^1 \leq H_{L^1+1}$. As shown in the proof of Lemma 1, $H_{L^1} < r^1$ is equivalent to $G_{L^1}^1 < r^1$, since $L^1 \leq L^*$. Let us now show that $r^1 \leq G_{L^1+1}^1$. If $f_{L^1+1}^* > 0$, this is equivalent to $r^1 \leq H_{L^1+1}$, as seen in the proof of Lemma 1. Therefore, we only need to consider the case where $f_{L^1+1}^* = 0$. In that case, $r^1 \leq G_{L^1+1}^1$ is equivalent to (33). Furthermore, (8) implies that $c_{L^1+1} - f_{L^1+1}^* = c_{L^1+1} \leq \sqrt{c_{L^1+1}/\lambda^*}$. Thus, to show (33) it suffices to show (34), which holds true. Hence, we have $G_{L^1}^1 < r^1 \leq G_{L^1+1}^1$.

Therefore, to establish that $\mathbf{f}^1 = \mathcal{N}^0(\mathbf{f}^0)$, it remains to be shown that $\frac{c_l^1}{(c_l^1 - f_l^1)^2} = \frac{c_m^1}{(c_m^1 - f_m^1)^2}$, for all $l, m \in \{1, \cdots, L^1\}$. Using (30) and (28), this is equivalent to showing $\frac{c_l}{(c_l - f_l^*)^2} = \frac{c_m}{(c_m - f_m^*)^2}$, for all $l, m \in \{1, \cdots, L^1\}$. This holds due to the optimality Conditions (4)–(5) for $(f_1^*, \cdots, f_L^*)$, since $L^1 \leq L^*$. This concludes the proof. □

Appendix B
Proofs of Results in Section VI

Proof of Lemma 2: Assume that there exists a maximally efficient strategy $\mathbf{f}^0$ of the leader, and let $\mathbf{f}^{-0} = \mathcal{N}^0(\mathbf{f}^0)$. Then, following precisely the proof of (21) in the single-follower case, one can show that for every $i \in \mathcal{I}$, the total flow sent by all other users on link l is

$$f_l^* - f_l^i = \begin{cases} c_l \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n}, & l = 1, \cdots, L^i \\ f_l^*, & l = L^i + 1, \cdots, L \end{cases} \tag{35}$$

and (26) follows. Similarly, it can be seen that, for any $i \in \mathcal{I}$, (14) holds, thus the threshold L^i is determined by (15). Moreover, using (35), one can show that (15) implies (25). Finally, using (26) and $\sum_{i \in \mathcal{I}_0} f_l^i = f_l^*$, $l \in \mathcal{L}$, (24) is immediate.

Suppose now that $\mathbf{f}^0$ given by (24) and (25) is an admissible strategy of the leader. If for all $i \in \mathcal{I}$, $\mathbf{f}^i$ is given by (26), it is easy to see that $\sum_{i \in \mathcal{I}_0} f_l^i = f_l^*$, $l \in \mathcal{L}$. Therefore, it suffices to show that $\mathbf{f}^{-0} = \mathcal{N}^0(\mathbf{f}^0)$, or equivalently, that $\mathbf{f}^i$ is the best reply of follower $i \in \mathcal{I}$ to the strategy profile $\mathbf{f}^{-i}$ of the other followers and the manager. It is easy to verify that for any link $l \in \mathcal{L}$ (35) holds. Observe that this is the maximally efficient strategy of the leader in a single-follower Stackelberg game where the follower has demand r^i and the demand of the leader is $(R - r^i)$, according to Theorem 1. Following precisely the proof of that theorem, one can show that $\mathbf{f}^i$ is indeed the best reply of user $i \in \mathcal{I}$ to $\mathbf{f}^{-i}$. □

Proof of Lemma 3: Suppose that $f_l^0 < 0$. Equation (24), then, gives

$$\sum_{i \in \mathcal{I}_l} \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} < (I_l - 1)\frac{f_l^*}{c_l} \leq (I_l - 1)\frac{f_{l-1}^*}{c_{l-1}} \tag{36}$$

since $f_l^*/c_l \leq f_{l-1}^*/c_{l-1}$, as implied by the optimality Conditions (4)–(5) for $(f_1^*, \cdots, f_L^*)$. If $\mathcal{I}_{l-1} = \mathcal{I}_l$, then $f_{l-1}^0 < 0$ is immediate from (36). Assume that $\mathcal{I}_{l-1} \backslash \mathcal{I}_l \neq \emptyset$. For all $i \in \mathcal{I}_{l-1} \setminus \mathcal{I}_l$, we have $L^i = l - 1$, and using inequality $H_{L^i} < r^i$, one can verify that $\frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} < \frac{f_{l-1}^*}{c_{l-1}}$. Summing this inequality over all $i \in \mathcal{I}_{l-1} \setminus \mathcal{I}_l$, and adding it to (36)

$$\sum_{i \in \mathcal{I}_{l-1}} \frac{\sum_{n=1}^{L^i} f_n^* - r^i}{\sum_{n=1}^{L^i} c_n} < [(I_l - 1) + (I_{l-1} - I_l)]\frac{f_{l-1}^*}{c_{l-1}}$$

thus $f_{l-1}^0 < 0$. □

Proof of Lemma 4: We give a sketch of the proof, that can be found in [25]. Using the methodology developed in [14], one can show that $\mathbf{f}^*$ is a continuous function of $r^0 \in [0, C - r]$. Similarly, $\mathbf{f}^0$ is a continuous function of $r^0 \in [0, C - r]$, except possibly of a finite number of points, $\alpha_1 < \cdots < \alpha_M$, where the threshold of at least one follower j is increased from L^j to $L^j + 1$. Therefore, we only need to establish continuity at α_m, $1 \leq m \leq M$. By its definition in (25), L^j is left-continuous at α_m and so is $\mathbf{f}^0$. Noting that at $r^0 = \alpha_m$, we have $r^j = H_{L^j+1}$, and using the continuity of $\mathbf{f}^*$, it can be easily seen that $\lim_{r^0 \downarrow \alpha_m} \mathbf{f}^0(r^0) = \lim_{r^0 \uparrow \alpha_m} \mathbf{f}^0(r^0)$. Thus $\mathbf{f}^0$ is continuous in $r^0 \in [0, C - r]$.

Since f_1^0 is continuous in $r^0 \in [0, C - r]$, in order to show that it is an increasing function, it suffices to show that it is increasing in every interval $[0, \alpha_1]$, $(\alpha_m, \alpha_{m+1}]$, $m = 1, \cdots, M - 1$, $(\alpha_M, C - r]$, where the threshold L^i of every follower i is constant. This can be easily verified observing that

$$\frac{f_1^0}{c_1} = \sum_{i=1}^{I-1} \left[\frac{\sum_{n=1}^{L^i} f_n^*}{\sum_{n=1}^{L^i} c_n} - \frac{f_1^*}{c_1} \right] + \frac{\sum_{n=1}^{L^I} f_n^*}{\sum_{n=1}^{L^I} c_n} - \sum_{i=1}^{I} \frac{r^i}{\sum_{n=1}^{L^i} c_n}.$$

□

Proof of Theorem 2: Recall that even if $\mathbf{f}^0$ is nonadmissible, it satisfies the demand constraint of the leader, according to (27). By virtue of Lemma 3, this implies that at $r^0 = 0$ we have $f_1^0 \leq 0$, since $f_1^0 > 0$ would imply $f_l^0 \geq 0$, for $l = 2, \cdots, L$, and the demand constraint of the leader would be violated. Suppose now that $r^0 = C - r$. Then $f_l^* = c_l$, $l \in \mathcal{L}$, and from (18) we have $H_l = 0$, for $l \in \mathcal{L}$, while

$H_{L+1} = R$. Thus, $L^i = L$ for every follower. Therefore, $f_1^0 = c_1 I(\sum_{n=1}^{L} f_n^* - r)/C - (I-1)f_1^* = c_1(1 - r/C) > 0$.

Since f_1^0 is continuous increasing in $[0, C-r]$, nonpositive at $r^0 = 0$ and positive at $r^0 = C - r$, there exists a unique $\underline{r}^0 \in [0, C-r)$, such that $f_1^0 = 0$ at $r^0 = \underline{r}^0$. Thus, $f_1^0 \geq 0$ if and only if $r^0 \in [\underline{r}^0, C-r)$, and the result follows. □

Proof of Proposition 2: By the definition of $\underline{r}^0$, it suffices to show that, with the demand r^0 of the leader fixed, f_1^0 is nonincreasing in the number of followers. Let $\mathbf{f}^0$ and $\hat{\mathbf{f}}^0$ be the strategy of the leader, given by (24) and (25), when there are I and $I+1$ followers, respectively. Note that in both cases the network optimum $(f_1^*, \cdots, f_L^*)$ is the same, since it depends on the total throughput demand $R = r^0 + r$, and not on the number of followers. Therefore, H_l is the same in both cases, for all $l \in \mathcal{L}$.

Since the followers are identical, their associated thresholds are equal, according to (25). Let L^1 and $\hat{L}^1$ be the thresholds when there are I and $I+1$ followers, respectively. In the former case, the demand of each follower is r/I and in the latter $r/(I+1)$. Therefore, (25) implies that $L^* \geq L^1 \geq \hat{L}^1$. In view of (24), to prove $f_1^0 \geq \hat{f}_1^0$, we have to show

$$\frac{I\sum_{n=1}^{L^1} f_n^* - r}{\sum_{n=1}^{L^1} c_n} - \frac{I\sum_{n=1}^{\hat{L}^1} f_n^* - r}{\sum_{n=1}^{\hat{L}^1} c_n} \geq \frac{\sum_{n=1}^{\hat{L}^1} f_n^*}{\sum_{n=1}^{\hat{L}^1} c_n} - \frac{f_1^*}{c_1}. \quad (37)$$

The expression on the right-hand side of (37) is nonpositive, since $f_1^*/c_1 \geq f_l^*/c_l$, for all $l \leq L^*$, as implied by the optimality conditions for $(f_1^*, \cdots, f_L^*)$. Therefore, it suffices to show that

$$\frac{I\sum_{n=1}^{L^1} f_n^* - r}{\sum_{n=1}^{L^1} c_n} \geq \frac{I\sum_{n=1}^{\hat{L}^1} f_n^* - r}{\sum_{n=1}^{\hat{L}^1} c_n}. \quad (38)$$

Since (38) holds trivially for $L^1 = \hat{L}^1$, we only need to consider the case $L^1 > \hat{L}^1$. Without loss of generality, assume that $\hat{L}^1 = L^1 - 1$. Then, (38) is equivalent to

$$\frac{r}{I} \geq \sum_{n=1}^{L^1-1} f_n^* - \frac{f_{L^1}^*}{c_{L^1}} \sum_{n=1}^{L^1-1} c_n = H_{L^1},$$

which is true, by the definition of the threshold L^1. □

Proof of Proposition 3: Suppose that $r^j \geq r^k$, and let $\mathbf{f}^0$ be the strategy of the leader given by (24). It suffices to show that if the demands of the followers become $r^j + \varepsilon$ and $r^k - \varepsilon$, $0 \leq \varepsilon \leq r^k$, and $\hat{\mathbf{f}}^0$ is the resulting strategy of the leader—according to (24)—then $\hat{f}_1^0 \geq f_1^0$. Since the total demand of the followers is fixed, the network optimum $(f_1^*, \cdots, f_L^*)$ and the threshold L^i of every follower $i \in \mathcal{I} \setminus \{j, k\}$ remain the same. Therefore, it suffices to show that

$$\phi(\varepsilon) \equiv \frac{\sum_{n=1}^{L^j} f_n^* - r^j - \varepsilon}{\sum_{n=1}^{L^j} c_n} + \frac{\sum_{n=1}^{L^k} f_n^* - r^k + \varepsilon}{\sum_{n=1}^{L^k} c_n} \quad (39)$$

is an nondecreasing function of $\varepsilon \in [0, r^k]$.

Note that L^j and L^k in (39) are also functions of ε. In particular, (25) implies that L^j is nondecreasing and L^k nonicreasing in ε. Then, it is easy to see that there exists a finite number of points $\alpha_1 < \cdots < \alpha_M$ in $(0, r^k)$, such that: 1) for all ε in the same interval $[0, \alpha_1]$, $(\alpha_m, \alpha_{m+1}]$, $m = 1, \cdots, M-1$, $(\alpha_M, r^k]$, both the thresholds L^j and L^k remain the same, and 2) at any point α_m, either L^j is increased, or L^k is decreased. Using the same technique as in the proof of Lemma 4, one can show that ϕ is a continuous function of $\varepsilon \in [0, r^k]$. Hence, to show that it is also nondecreasing, it suffices to show that it is nondecreasing in every $(\alpha_m, \alpha_{m+1}]$ interval, where the thresholds L^j and L^k are fixed. But this is immediate from (39), since $L^j \geq L^k$ implies that $\varepsilon(1/\sum_{n=1}^{L^k} c_n - 1/\sum_{n=1}^{L^j} c_n)$ is nondecreasing in ε. □

REFERENCES

[1] R. B. Myerson, *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard Univ. Press, 1991.

[2] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1992.

[3] M.-T. T. Hsiao and A. A. Lazar, "Optimal decentralized flow control of markovian queueing networks with multiple controllers," *Performance Evaluation*, vol. 13, no. 3, pp. 181–204, 1991.

[4] Z. Zhang and C. Douligeris, "Convergence of synchronous and asynchronous greedy algorithms in a multiclass telecommunications environment," *IEEE Trans. Commun.*, vol. 40, pp. 1277–1281, Aug. 1992.

[5] E. Altman, "Flow control using the theory of zero-sum Markov games," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 814–818, Apr. 1994.

[6] Y. A. Korilis and A. A. Lazar, "On the existence of equilibria in noncooperative optimal flow control," *J. ACM*, vol. 42, pp. 584–613, May 1995.

[7] A. A. Economides and J. A. Silvester, "Multiobjective routing in integrated services networks: A game theory approach," in *Proc. INFOCOM'91*, 1991, pp. 1220–1225.

[8] A. Orda, R. Rom, and N. Shimkin, "Competitive routing in multiuser communication networks," *IEEE/ACM Trans. Networking*, vol. 1, pp. 510–521, Oct. 1993.

[9] A. A. Lazar, A. Orda, and D. E. Pendarakis, "Virtual path bandwidth allocation in multiuser networks," in *Proc. IEEE INFOCOM'95*, Boston, MA, Apr. 1995, pp. 312–320.

[10] P. Dubey, "Inefficiency of Nash equilibria," *Math. Oper. Res.*, vol. 11, pp. 1–8, Feb. 1986.

[11] R. Cocchi, S. Shenker, D. Estrin, and L. Zhang, "Pricing in computer networks: Motivation, formulation and example," *IEEE/ACM Trans. Networking*, vol. 1, pp. 614–627, Dec. 1993.

[12] J. K. MacKie-Mason and H. R. Varian, "Pricing cognestible network resources," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1141–1149, Sept. 1995.

[13] S. J. Shenker, "Making greed work in networks: A game-theoretic analysis of switch service disciplines," *IEEE/ACM Trans. Networking*, vol. 3, pp. 819–831, Dec. 1995.

[14] Y. A. Korilis, A. A. Lazar, and A. Orda, "Capacity allocation under noncooperative routing," *IEEE Trans. Automat. Contr.*, to appear.

[15] Information Sciences Institute, "Internet protocol," Univ. Southern California, Marina del Rey, CA, Tech. Rep. RFC 791, Sept. 1981.

[16] S. Deering and R. Hinden, "Internet protocol version 6 specification," Internet Draft, IETF, Mar. 1995.

[17] J. J. Garrahan, P. A. Russo, K. Kitami, and R. Kung, "Intelligent network overview," *IEEE Commun. Mag.*, vol. 31, pp. 30–36, Mar. 1993.

[18] P. A. Russo, K. Bechard, E. Brooks, R. L. Corn, R. Gove, W. L. Honig, and J. Young, "IN rollout in the United States," *IEEE Commun. Mag.*, vol. 31, pp. 56–63, Mar. 1993.

[19] C. Douligeris and R. Mazumdar, "Multilevel flow control of queues," in *Proc. Johns Hopkins Conference on Information Sciences*, Baltimore, MD, Mar. 1989.

[20] A. A. Economides and J. A. Silvester, "Priority load sharing: An approach using Stackelberg games," in *Proc. 28th Annu. Allerton Conf. Commun., Control, Computing*, 1990.

[21] M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*. Reading, MA: Addison-Wesley, 1987.

[22] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[23] D. G. Luenberger, *Linear and Nonlinear Programming*. Reading, MA: Addison Wesley, 1984.

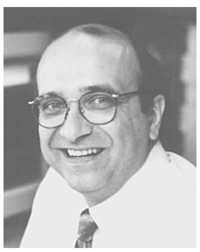
[24] S. Shenker, "Service models and pricing policies for an integrated services internet," in *Public Access to the Internet*, B. Kahin and J. Keller, Eds. Cambridge, MA: MIT Press, 1993, pp. 315–337.

[25] Y. A. Korilis, A. A. Lazar, and A. Orda, "Achieving network optima using Stackelberg routing strategies," Center for Telecommunications Research, Columbia University, NY, CTR Tech. Rep. 384-94-31, 1994.

Yannis A. Korilis (S'89–M'95) was born in Thessaloniki, Greece, in 1966. He received the Diploma from the National Technical University of Athens, Greece, in 1989, and the M.S., M.Phil., and Ph.D. degrees from Columbia University, NY, in 1990, 1992, and 1995, respectively, all in electrical engineering.

Since 1995, he has been a Member of the Technical Staff, Bell Labs/Lucent Technologies, Holmdel, NJ. His research interests include control and management of multimedia networks, pricing of multimedia services, economic and game theoretic models for networking, distributed algorithms. During the spring of 1996, he held an Adjunct Assistant Professor position at Columbia University, NY. During the summer of 1993, he worked as Senior Staff Technologist at Bellcore, Morristown, NJ.

Dr. Korilis has received various academic excellence awards from the National Scholarship Foundation and the Technical Chamber of Greece. He has been a winner in the Interschool Fellowship Competition at Columbia University, for four consecutive years. He has also received the E. I. Jury Award for excellence in the area of Systems, Communications and Signal Processing. He is a member of the ACM and INFORMS.



Aurel A. Lazar (S'77–M'80–SM'90–F'93) is a Professor of Electrical Engineering at Columbia University, NY. His research interests span both theoretical and experimental studies of telecommunication networks and multimedia systems. The theoretical research he conducted during the 1980's pertains to the modeling, analysis and control of broadband networks. He formulated optimal flow and admission control problems and, by building upon the theory of point processes, derived control laws for Markovian queueing network models in single control as well as game theoretic settings. He was the chief architect of two experimental networks, generically called MAGNET. This work introduced traffic classes with explicit quality-of-service constraints to broadband switching and led to the concepts of schedulable, admissible load and contract regions in real-time control of broadband networks. His current research in broadband networking with quality-of-service guarantees focusses on modeling of video streams and analysing their multiplexing behavior, with emphasis on multiple time scales and subexponentiality. He is also leading investigations into multimedia networking architectures supporting interoperable exchange mechanisms for interactive and on demand multimedia applications with quality of service requirements. The main focus of this work is on building an open signalling infrastructure (also known as the xbind binding architecture) that enables the rapid creation, deployment and management of multimedia services.

In the early 1990's, his research efforts shifted to the foundations of the control, management and telemedia architecture of future multimedia networks. His involvement with gigabit networking research lead to the first fully operational service management system on ATM-based broadband networks. The system was implemented on top of AT&T's XUNET III gigabit platform spanning the continental US. His management and control research pioneered the application of virtual reality to the management of ATM-based broadband networks. He was instrumental in establishing the OPENSIG (http://www.ctr.columbia.edu/opensig) international working group with the goal of exploring network programability and next generation signalling technology.

Ariel Orda (S'84–M'92) was born in Argentina in 1961. He received the B.Sc. (Summa Cum Laude), M.Sc. and D.Sc. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1983, 1985, and 1991, respectively.

From 1983 to 1985, he was a Teaching Assistant at the Technion–Israel Institute of Technology, Haifa, Israel. From 1985 to 1990, he served in the IDF as a Research Engineer. From 1990 to 1991, he was a Teaching Instructor at the Technion. In 1991, he joined the Faculty of Electrical Engineering at the Technion, as a Lecturer. During the academic year 1993–1994, he was a Visiting Scientist with the Center for Telecommunications Research, Columbia University, NY. Since 1994, he has been a Senior Lecturer at the Technion. From 1991 to 1996, he held consulting positions in AT&T Bell Laboratories and in the Israeli industry. During the summers of 1995 and 1996, he was an Academic Visitor at IBM T. J. Watson Research Center, Yorktown Heights, NY. His current research interests include the control and management of broadband networks, the application of game theory to computer networking, QoS routing, and on-line network algorithms.

Dr. Orda received the award of the Chief Scientist in the Ministry of Communication and a Gutwirth Award for Outstanding Distinction in 1991. In 1993, he received the research award of the Association of Computer and Electronic Industries in Israel.