Cross-Layer Optimization in TCP/IP Networks

Jiantao Wang, *Student Member, IEEE*, Lun Li, *Student Member, IEEE*, Steven H. Low, *Senior Member, IEEE*, and John C. Doyle

***Abstract*—TCP-AQM can be interpreted as distributed primal-dual algorithms to maximize aggregate utility over source rates. We show that an equilibrium of TCP/IP, if exists, maximizes aggregate utility over both source rates and routes, provided congestion prices are used as link costs. An equilibrium exists if and only if this utility maximization problem and its Lagrangian dual have no duality gap. In this case, TCP/IP incurs no penalty in not splitting traffic across multiple paths. Such an equilibrium, however, can be unstable. It can be stabilized by adding a static component to link cost, but at the expense of a reduced utility in equilibrium. If link capacities are optimally provisioned, however, pure static routing, which is necessarily stable, is sufficient to maximize utility. Moreover single-path routing again achieves the same utility as multipath routing at optimality.**

***Index Terms*—Congestion control, routing, TCP-AQM/IP, utility maximization.**

I. Introduction

RECENT studies have shown that any TCP congestion control algorithm can be interpreted as carrying out a distributed primal-dual algorithm over the Internet to maximize aggregate utility, and a user's utility function is (often implicitly) defined by its TCP algorithm, see, e.g., [15], [19], [24], [23], [21], [16], [18] for unicast and [13], [7] for multicast. See also [20], [14], [27] for recent surveys and further references. All of these works assume that routing is given and fixed at the timescale of interest, and TCP, together with active queue management (AQM), attempt to maximize aggregate utility over source rates. In this paper, we study cross-layer utility maximization at the timescale of route changes.

We focus on the situation where a single minimum-cost route (shortest path) is selected for each source-destination pair (Section II). This models IP routing in the current Internet within an Autonomous Systems using common routing protocols such as OSPF [25][1] or RIP [11]. Routing is typically updated at a much slower timescale than TCP-AQM. We model this by assuming that TCP and AQM converge instantly to equilibrium after each route update to produce source rates and "congestion prices" for that update period. These congestion prices may represent delays or loss probabilities across network links. They determine the next routing update in the case of dynamic routing, similar to the system analyzed in [10]. Thus TCP-AQM/IP form a feedback system where routing interacts with congestion control in an iterative process. We are interested in the equilibrium and stability properties of this iterative process. To simplify notation, we will henceforth use TCP-AQM/IP and TCP/IP interchangeably.

A. Summary

Here are our main results. In the case of pure dynamic routing, i.e., when link costs are the congestion prices generated by TCP-AQM, it turns out that we can interpret TCP/IP as a distributed primal-dual algorithm to maximize aggregate utility over *both* source rates (by TCP-AQM) and routes (by IP), in the following sense (Section III). We consider the problem, and its Lagrangian dual, of maximizing utility over source rates and over routing that use only a *single* path for each source-destination pair. Unlike the TCP-AQM problem or the multipath routing problem (see below) that are convex optimizations with no duality gap, the single-path TCP/IP problem is nonconvex and generally has a duality gap. An equilibrium of the TCP/IP system exists if and only if this problem has no duality gap. In this case, a TCP/IP equilibrium solves both the primal and the dual problem. Moreover, it incurs no penalty for not splitting traffic across multiple paths: optimal single-path routing achieves the same aggregate utility as optimal multipath routing. Multipath routing can achieve a strictly higher utility only when there is a duality gap between the single-path primal and dual problems, but in this case, the TCP/IP iteration does not even have an equilibrium, let alone solving the utility maximization problem.

Even when the single-path problem has no duality gap and TCP/IP has an equilibrium, the equilibrium is generally unstable under pure dynamic routing. It can be stabilized by adding a sufficiently large static component to the definition of link cost. The existence and characterization of TCP/IP equilibrium when the link costs are not pure congestion prices, however, are open problems. To proceed, we specialize to a ring network with a common destination and demonstrate an inevitable tradeoff between utility maximization and routing stability (Section IV). Specifically, we show that the TCP/IP system over the special ring network is indeed unstable when link costs are pure prices. It can be stabilized by adding a static component to the link cost, but at the expense of a reduced utility in equilibrium. The loss in utility increases with the weight on the static component. Hence, while stability requires a small weight on prices, utility maximization favors a large weight. We present numerical results to validate these qualitative conclusions in a general network topology. They also suggest that routing instability can reduce

Manuscript received December 30, 2002; revised June 12, 2004; approved by IEEE/ACM Transactions on Networking Editor R. Srikant. This work is part of the FAST project funded by the National Science Foundation, the Caltech Lee Center for Advanced Networking, the Army Research Office (ARO), the Air Force Office of Scientific Research (AFOSR), the Defense Advanced Research Projects Agency (DARPA), and Cisco. Partial and preliminary results appeared in the Proceedings of IEEE INFOCOM, San Francisco, CA, April 2003.

The authors are with the California Institute of Technology, Pasadena, CA 91125 USA (e-mail: jiantao@cds.caltech.edu; lun@cds.caltech.edu; slow@caltech.edu; doyle@cds.caltech.edu).

Digital Object Identifier 10.1109/TNET.2005.850219

[1]Even though OSPF implements a shortest-path algorithm, it allows multiple equal-cost paths to be utilized. Our model ignores this feature.

aggregate utility to less than that achievable by (the necessarily stable) pure static routing.

Indeed we show that if the link capacities are optimally provisioned, then *pure static* routing is enough to maximize utility even for general networks (Section V). Moreover, it is optimal within the class of multipath routing: again, there is no penalty at optimality in not splitting traffic across multiple paths.

Finally, we discuss some implications and limitations of these results (Section VI).

B. Related Work

Our goal is to understand equilibrium and stability issues in cross-layer optimization of TCP/IP networks. Another approach to joint routing and congestion control is to allow multipath routing, i.e., a source can transmit its data along multiple paths to its destination. In this formulation, a source's decision is decomposed into two—how much traffic to send (congestion control) and how to distributed it over the available paths (multi-path routing or load balancing)—in order to maximize aggregate utility. This has been analyzed in, e.g., [8], [15], [12]. The general intuition is that, for each source-destination pair, only paths with the minimum, and hence equal, "congestion prices" will be used and this minimum price determines the total source rate as in the single-path case. In contrast to TCP/IP networks, this formulation assumes that both decisions operate on the same timescale. It however provides an upper bound to the problem TCP/IP attempts to solve (see Section III-A).

The multipath problem is equivalent to multicommodity flow problem which has been extensively studied; see, e.g., [1], [5]. The standard formulation is to maximize aggregate throughput, corresponding to a common and linear utility function. It is then a linear program and therefore can be solved in polynomial time, though there are combinatorial algorithms for this class of problems that are more efficient. Recently, fast approximation algorithms and their competitive ratios have been developed for network flow, and other, problems, e.g., [26], [10], [2]. Since multipath problem upper bounds the TCP/IP problem, the work on network flow problems provides insight on the performance limit of TCP/IP. There are however differences. First, our single-path routing problem is NP-hard (see Section III-A) and generally has a duality gap, whereas the network flow problem is generally a linear program that is in P and has no duality gap. Second, the utility functions that correspond to common TCP algorithms are strictly concave whereas they are typically linear, in fact, identity, functions in network flow problems. Third, the algorithms developed for network flow problems are typically centralized and therefore cannot model TCP/IP iterations or be carried out in a large network where they must be decentralized.

Instability of single-path routing is not surprising as it is well-known that stability generally requires that the relative weight on the dynamic (traffic-sensitive) component of the link cost be small. Indeed, our conclusions are similar to those reached in [4], [17] that study the same ring network for routing stability using different link costs. Here, since the dynamic component is the dual-optimal price for the utility maximization problem computed by TCP-AQM, this implies a tradeoff between routing stability and utility maximization.

II. Model

In general, we use small letters to denote vectors, e.g., x with x_i as its ith component; capital letters to denote matrices, e.g., H, W, R, or constants, e.g., L, N, K^i; and script letters to denote sets of vectors or matrices, e.g., $\mathcal{W}_s, \mathcal{W}_m, \mathcal{R}_s, \mathcal{R}_m$. Superscript is used to denote vectors, matrices, or constants pertaining to source i, e.g., y^i, w^i, H^i, K^i.

A. Network

A network is modeled as a set of L uni-directional links with finite capacities $c = (c_l, l = 1, \ldots, L)$, shared by a set of N source-destination pairs, indexed by i (we will also refer to the pair simply as "source i"). There are K^i acyclic paths for source i represented by a $L \times K^i$ 0-1 matrix H^i where

$$H^i_{lj} = \begin{cases} 1, & \text{if path } j \text{ of source } i \text{ uses link } l \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathcal{H}^i$ be the set of all columns of H^i that represents all the available paths to i under single-path routing. Define the $L \times K$ matrix H as

$$H = [H^1 \quad \cdots \quad H^N]$$

where $K := \sum_i K^i$. H defines the topology of the network.

Let w^i be a $K^i \times 1$ vector where the jth entry represents the fraction of i's flow on its jth path such that

$$w^i_j \geq 0 \quad \forall j, \quad \text{and} \quad \mathbf{1}^T w^i = 1$$

where $\mathbf{1}$ is a vector of an appropriate dimension with the value 1 in every entry. We require $w^i_j \in \{0, 1\}$ for single path routing, and allow $w^i_j \in [0, 1]$ for multipath routing. Collect the vectors $w^i, i = 1, \ldots, N$, into a $K \times N$ block-diagonal matrix W. Let $\mathcal{W}_s$ be the set of all such matrices corresponding to single path routing defined as

$$\{W \mid W = \text{diag}(w^1, \ldots, w^N) \in \{0, 1\}^{K \times N}, \mathbf{1}^T w^i = 1\}.$$

Define the corresponding set $\mathcal{W}_m$ for multipath routing as

$$\{W \mid W = \text{diag}(w^1, \ldots, w^N) \in [0, 1]^{K \times N}, \mathbf{1}^T w^i = 1\}. \quad (1)$$

As mentioned above, H defines the set of acyclic paths available to each source, and represents the network topology. W defines how the sources load balance across these paths. Their product defines a $L \times N$ routing matrix $R = HW$ that specifies the fraction of i's flow at each link l. The set of all single-path routing matrices is

$$\mathcal{R}_s = \{R \mid R = HW, W \in \mathcal{W}_s\} \quad (2)$$

and the set of all multipath routing matrices is

$$\mathcal{R}_m = \{R \mid R = HW, W \in \mathcal{W}_m\}. \quad (3)$$

The difference between single-path routing and multipath routing is the integer constraint on W and R. A single-path routing matrix in $\mathcal{R}_s$ is an 0-1 matrix:

$$R_{li} = \begin{cases} 1, & \text{if link } l \text{ is in a path of source } i \\ 0, & \text{otherwise.} \end{cases}$$

A multipath routing matrix in $\mathcal{R}_m$ is one whose entries are in the range [0, 1]:

$$R_{li} \begin{cases} >0, & \text{if link } l \text{ is in a path of source } i \\ =0, & \text{otherwise.} \end{cases}$$

The path of source i is denoted by $r^i = [R_{1i} \ \cdots \ R_{Li}]^T$, the ith column of the routing matrix R.

B. TCP-AQM/IP

We consider the situation where TCP-AQM operates at a faster timescale than routing updates. We assume a *single* path is selected for each source-destination pair that minimizes the sum of the link costs in the path, for some appropriate definition of link cost. In particular, traffic is not split across multiple paths from the source to the destination even if they are available. This models, e.g., IP routing within an Autonomous System. We focus on the timescale of the route changes, and assume TCP-AQM is stable and converges instantly to equilibrium after a route change. As in [18], we will interpret the equilibria of various TCP and AQM algorithms as solutions of a utility maximization problem defined in [15]. Different TCP algorithms solve the same prototypical problem (4) with different utility functions [18], [21].

Specifically, suppose each source i has a utility function $U_i(x_i)$ as a function of its total transmission rate x_i. We assume U_i is strictly concave increasing (which is the case for common TCP algorithms [18]). Let $R(t) \in \mathcal{R}_s$ be the (single-path) routing in period t. Let the equilibrium rates $x(t) = x(R(t))$ and prices $p(t) = p(R(t))$ generated by TCP-AQM in period t, respectively, be the optimal solutions of the constrained maximization problem

$$\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{s.t. } R(t)x \leq c \tag{4}$$

and its Lagrangian dual

$$\min_{p \geq 0} \sum_i \max_{x_i \geq 0} \left(U_i(x_i) - x_i \sum_l R_{li}(t) p_l \right) + \sum_l c_l p_l. \tag{5}$$

The prices $p_l(t), l = 1, \ldots, L$, are measures of congestion, such as queueing delays or loss probabilities [18], [21]. We assume the link costs in period t are

$$d_l(t) = a p_l(t) + b \tau_l \tag{6}$$

where $a \geq 0, b \geq 0$, and $\tau_l > 0$ are constants. Based on these costs, each source computes its new route $r^i(t+1) \in \mathcal{H}^i$ individually that minimizes the total cost on its path:

$$r^i(t+1) = \arg \min_{r^i \in \mathcal{H}^i} \sum_l d_l(t) r^i_l. \tag{7}$$

In (6), τ_l are traffic insensitive components of the link cost $d_l(t)$, e.g., $\tau_l = 1/c_l$. If τ_l represent the fixed propagation delays across links l and $p_l(t)$ the queueing delays at links l, then $d_l(t)$ represent total delays across links l. The protocol parameters a and b determine the responsiveness of routing to network traffic: $a = 0$ corresponds to static routing, $b = 0$ corresponds to purely dynamic routing, and the larger the ratio of a/b, the more responsive routing is to network traffic. As we will see below, they determine whether an equilibrium exists, whether it is stable, and the achievable utility at equilibrium.

An equivalent way to specify the TCP-AQM/IP system as a dynamical system, at the timescale of route changes, is to replace (4)–(5) by their optimality conditions. The routing is updated according to

$$r^i(t+1) = \arg \min_{r^i \in \mathcal{H}^i} \sum_l (a p_l(t) + b \tau_l) r^i_l, \quad \text{for all } i \tag{8}$$

where $p(t)$ and $x(t)$ are given by

$$\sum_l R_{li}(t) p_l(t) = U_i'(x_i(t)) \quad \text{for all } i \tag{9}$$

$$\sum_i R_{li}(t) x_i(t) \begin{cases} \leq c_l & \text{if } p_l(t) \geq 0 \\ = c_l & \text{if } p_l(t) > 0 \end{cases} \quad \text{for all } l \tag{10}$$

$$x(t) \geq 0, \quad p(t) \geq 0. \tag{11}$$

This set of equations describe how the routing $R(t)$, rates $x(t)$, and prices $p(t)$ evolve. Note that $x(t)$ and $p(t)$ depend only on $R(t)$ through (9)–(11), implicitly assuming that TCP-AQM converges instantly to an equilibrium given the new routing $R(t)$.

We say that (R^*, x^*, p^*) is an *equilibrium of TCP/IP* if it is a fixed point of (4)–(7), or equivalently, (8)–(11), i.e., starting from routing R^* and associated (x^*, p^*), the above iterations yield (R^*, x^*, p^*) in the subsequent periods.

III. Equilibrium of TCP/IP

In this section, we study the condition under which TCP/IP as modeled by (4)–(7) or (8)–(11) has an equilibrium, and characterize the equilibrium as the optimal solution of an optimization problem. Since (8)–(11) is a system of mixed integer nonlinear inequalities, characterization of its equilibrium, even the basic question of existence and uniqueness, is in general difficult to determine. The case of pure dynamic routing, $a > 0$ and $b = 0$, is the simplest and most instructive.

A. Pure Dynamic Routing: Main Results

In the following, we completely characterize the case of pure dynamic routing, $a > 0$ and $b = 0$. Without loss of generality, we set $a = 1$ in (7) and (8) when $b = 0$.

Consider the problem

$$\max_{R \in \mathcal{R}_s} \max_{x \geq 0} \sum_i U_i(x_i) \quad \text{s.t. } Rx \leq c \tag{12}$$

and its Lagrangian dual

$$\min_{p \geq 0} \sum_i \max_{x_i \geq 0} \left(U_i(x_i) - x_i \min_{r^i \in \mathcal{H}^i} \sum_l R_{li} p_l \right) + \sum_l c_l p_l \tag{13}$$

where r^i is the ith column of R with $r^i_l = R_{li}$. While (4)–(5) maximize utility over source rates only, problem (12) maximizes utility over both rates and routes. While (4) is a convex program without duality gap, problem (12) is nonconvex because the variable R is discrete, and generally has a duality gap.[2] The interesting feature of the dual problem (13) is that the maximization over R takes the form of minimum-cost routing with prices p generated by TCP-AQM as link costs. This suggests that TCP/IP might turn out to be a distributed algorithm that attempts to maximize utility, with proper choice of link costs. This is indeed true—when equilibrium of TCP/IP exists.

Theorem 1: Suppose $a = 1, b = 0$.

1) An equilibrium (R^*, x^*, p^*) of TCP/IP exists if and only if there is no duality gap between (12) and (13).
2) In this case, the equilibrium (R^*, x^*, p^*) is a solution of (12) and (13).

Hence, one can regard the layering of TCP and IP as a decomposition of the utility maximization problem over source rates and routes into a distributed and decentralized algorithm, carried out on two different timescales, in the sense that an equilibrium of the TCP/IP iteration (8)–(11), if it exists, solves (12) and (13). An equilibrium may not exist. Even if it does, it may not be stable—an issue we address in Section IV.

Example 1: Duality Gap: A simple example where there is a duality gap and equilibrium of TCP/IP does not exist consists of a single source-destination pair connected by two parallel links each of capacity 1, as shown in Fig. 2 (take $N = 1$). Clearly, under pure dynamic single-path routing, equilibrium of TCP/IP does not exist, because the TCP/IP iteration (8)–(11) will choose one of the two routes in each period to carry all traffic. TCP-AQM will generate positive price for the chosen route and zero price for the other route, so that in the next period, the other route will be selected, and the cycle repeats. The proof that there is a duality gap between the primal problem (12) and the dual problem (13) is given in Appendix VII.A (take $N = 1$). Intuitively, either path is optimal (both for primal and for dual problem). For the primal problem the optimal rate is $x^* = 1$, constrained by link capacity, whereas for the dual problem, the optimal rate is $x^* = 2$, primal infeasible. Hence the primal optimal value is $U(1)$, strictly less than the dual optimal value of $U(2)$. □

The duality gap is a measure of "cost of not splitting".[3] To elaborate, define the Lagrangian [3], [22]

$$L(R, x, p) = \sum_i \left(U_i(x_i) - x_i \sum_l R_{li} p_l \right) + \sum_l c_l p_l.$$

The primal (12) and dual (13) can then be expressed respectively as

$$V_{\rm sp} = \max_{R \in \mathcal{R}_s, x \geq 0} \min_{p \geq 0} L(R, x, p)$$
$$V_{\rm sd} = \min_{p \geq 0} \max_{R \in \mathcal{R}_s, x \geq 0} L(R, x, p).$$

[2]The nonlinear constraint $Rx \leq c$ can be converted into a linear constraint—see Proof of Theorem 2—so integer constraint on R is the real source of difficulty.

[3]A term apparently coined by Bruce Hayek.

If we allow sources to distribute their traffic among multiple paths available to them, then the corresponding problems for multipath routing are

$$V_{\rm mp} = \max_{R \in \mathcal{R}_m, x \geq 0} \min_{p \geq 0} L(R, x, p)$$
$$V_{\rm md} = \min_{p \geq 0} \max_{R \in \mathcal{R}_m, x \geq 0} L(R, x, p). \quad (14)$$

Since $\mathcal{R}_s \subseteq \mathcal{R}_m, V_{\rm sp} \leq V_{\rm mp}$. The next result clarifies the relation among these four problems.

Theorem 2:

$$V_{\rm sp} \leq V_{\rm sd} = V_{\rm mp} = V_{\rm md}.$$

According to Theorem 1, TCP/IP has an equilibrium exactly when there is no duality gap in the single-path utility maximization, i.e., when $V_{\rm sp} = V_{\rm sd}$. Theorem 2 then says that in this case, there is no penalty in not splitting the traffic, i.e., single-path routing performs as well as multipath routing, $V_{\rm sp} = V_{\rm mp}$. Multipath routing achieves a strictly higher utility $V_{\rm mp}$ precisely when TCP/IP has no equilibrium, in which case the TCP/IP iteration (8)–(11) cannot converge, let alone solving the single-path utility maximization problem (12) or (13). In this case the problem (12) and its dual (13) do not characterize TCP/IP, but their gap measures the loss in utility in restricting routing to single-path and is of independent interest.

Even though minimum-cost routing is polynomial, it is shown in [28] that single-path utility maximization is NP-hard. This is not surprising since, e.g., a related problem on load balancing on a ring has been proved to be NP-hard in [6].

Theorem 3: The primal problem (12) is NP-hard.

Theorem 3 shows that the general problem (12) is NP-hard, by reducing all instances of the integer partition problem to some instances of the primal problem (12). Theorem 2 however implies that the sub-class of the utility maximization problems with no duality gap are in P, since they are equivalent to multipath problems which are concave programs and hence polynomial-time solvable. It is a common phenomenon for sub-classes of NP-hard problems to have polynomial-time algorithms (e.g., satisfiability is NP-hard, and yet 2-SAT is in P). Informally, the hard problems are those with nonzero duality gap.

B. Pure Dynamic Routing: Proofs

In this subsection, we present Proofs for Theorems 1–3. We will first prove Theorem 2. Then we show that an equilibrium of TCP/IP must solve the dual problem (13). This together with Theorem 2 imply Theorem 1. Finally, we present a Proof for Theorem 3.

Proof of Theorem 2: The inequality follows from the weak duality theorem [3]. We now prove $V_{\rm sd} = V_{\rm md}$ and $V_{\rm mp} = V_{\rm md}$.

We have

$$V_{\rm sd} = \min_{p \geq 0} \max_{R \in \mathcal{R}_s, x \geq 0} \left(\sum_i U_i(x_i) - p^T R x \right) + p^T c$$
$$= \min_{p \geq 0} \max_{x \geq 0} \left(\sum_i U_i(x_i) - \min_{W \in \mathcal{W}_s} p^T H W x \right) + p^T c$$

where $R = HW$ with $W \in \mathcal{W}_s$ from (2). Similarly, from (3) we have

$$V_{\rm md} = \min_{p\geq 0}\max_{x\geq 0}\left(\sum_i U_i(x_i) - \min_{W\in\mathcal{W}_m} p^T HWx\right) + p^T c.$$

Define functions $f_s(x,p)$ and $f_m(x,p)$ as:

$$f_s(x,p) := \min_{W\in\mathcal{W}_s} p^T HWx$$
$$f_m(x,p) := \min_{W\in\mathcal{W}_m} p^T HWx.$$

In order to show that $V_{\rm sd} = V_{\rm md}$, we only need to show that $f_s(x,p) = f_m(x,p)$. Clearly $f_s(x,p) \geq f_m(x,p)$ since $\mathcal{W}_s \subseteq \mathcal{W}_m$. From (1), noting that $W = \mathrm{diag}(w^i)$, we have

$$\begin{aligned} f_m(x,p) &= \min_W p^T HWx \\ &\text{s.t.}\quad \mathbf{1}^T w^i = 1,\quad 0 \leq w_j^i \leq 1. \end{aligned}$$

Since this is a linear program for the given x and p, at least one of the optimal points lies on the boundary, i.e., $w_j^i = 0$ or 1 for all i and j, and hence is in $\mathcal{W}_s \subseteq \mathcal{W}_m$. Such a point solves both $f_s(x,p)$ and $f_m(x,p)$, i.e., $f_s(x,p) = f_m(x,p)$.

To show $V_{\rm md} = V_{\rm mp}$, transform $V_{\rm mp}$ into a convex optimization with linear constraints, which hence has no duality gap; see, e.g., [3]. Now, $V_{\rm mp}$ is equivalent to the problem

$$\max_{R\in\mathcal{R}_m, x\geq 0} \sum_i U_i(x_i) \quad \text{s.t. } Rx \leq c. \tag{15}$$

Note that this is not a convex program since the feasible set specified by $Rx \leq c$ is generally not convex. Define the $K_i \times 1$ vectors y^i in terms of the scalar x_i and the $K_i \times 1$ vectors w^i as the new variables

$$y^i = x_i w^i. \tag{16}$$

The mapping from (x_i, w^i) to y^i is one-to-one: the inverse of (16) is $x_i = \mathbf{1}^T y^i$ and $w^i = y^i / x_i$.

Now change the variables in (15) and (14) from (W,x) to y—substituting $x_i = \mathbf{1}^T y^i$ and $Rx = HWx = Hy$ into (15) and (14). We obtain an equivalent problem

$$\max_{y\geq 0} \sum_i U_i(\mathbf{1}^T y^i) \quad \text{s.t. } Hy \leq c.$$

and its Lagrangian dual. This is a convex program with linear constraint and hence has no duality gap. This proves $V_{\rm mp} = V_{\rm md}$. □

Proof of Theorem 1: It is easy to show that optimal solutions exist for both the primal problem (12) and its dual (13), so the issue is whether there is a duality gap. We will prove the theorem in two steps. First, given an equilibrium $(\tilde{R}, \tilde{x}, \tilde{p})$ of TCP/IP, we will show that it solves both the primal (12) and the dual (13), and hence there is no duality gap. Then, given a solution (R^*, x^*, p^*) of the primal and the dual problems, we will show that it is an equilibrium of TCP/IP.

Step 1: Necessity. Let $(\tilde{R}, \tilde{x}, \tilde{p})$ be an equilibrium of TCP/IP, i.e., a fixed point of (4)–(7) with $a = 1, b = 0$. Then

$$\tilde{p}^T \tilde{r}^i = \min_{r^i\in\mathcal{H}^i} \tilde{p}^T r^i \quad \text{for all } i, \tag{17}$$

$$(\tilde{p}, \tilde{x}) = \arg\min_{p\geq 0}\max_{x\geq 0}\left(\sum_i U(x_i) - p^T \tilde{R} x\right) + p^T c \tag{18}$$

where, again, r^i are the ith columns of routing matrix $R \in \mathcal{R}_s$.[4] We will show that $(\tilde{R}, \tilde{x}, \tilde{p})$ solves the dual problem (13). Then, since the dual problem (13) upper bounds the primal problem (12) (Theorem 2), and $\tilde{R} \in \mathcal{R}_s$ is a single-path routing and hence primal feasible, $(\tilde{R}, \tilde{x}, \tilde{p})$ also solves the primal (12).

To show that $(\tilde{R}, \tilde{x}, \tilde{p})$ solves the dual problem, we use the fact that the dual problem has an optimal solution, denoted by (R^*, x^*, p^*), and show that both achieve the same dual objective value, i.e., $L(\tilde{R}, \tilde{x}, \tilde{p}) = L(R^*, x^*, p^*)$. Now

$$(p^*, x^*, R^*) = \arg\min_{p\geq 0}\max_{x\geq 0}\left(\sum_i U(x_i) - \min_{R\in\mathcal{R}_s} p^T Rx + p^T c\right). \tag{19}$$

Let

$$f(p) := \max_{x\geq 0}\left(\sum_i U(x_i) - p^T \tilde{R} x\right) + p^T c,$$
$$g(p) := \max_{x\geq 0}\left(\sum_i U(x_i) - \min_{R\in\mathcal{R}_s} p^T Rx\right) + p^T c.$$

Then (18) implies $f(\tilde{p}) = \min_{p\geq 0} f(p)$ and (19) implies $g(p^*) = \min_{p\geq 0} g(p)$. Since $\tilde{R} \in \mathcal{R}_s$, we have

$$f(p) \leq g(p) \quad \text{for all } p \geq 0$$

and hence

$$f(\tilde{p}) = \min_{p\geq 0} f(p) \leq \min_{p\geq 0} g(p) = g(p^*).$$

On the other hand

$$\begin{aligned} f(\tilde{p}) &= \max_{x\geq 0} \sum_i U(x_i) - \tilde{p}^T \tilde{R} x + \tilde{p}^T c \\ &= \max_{x\geq 0} \sum_i U(x_i) - \sum_i x_i \left(\tilde{p}^T \tilde{r}^i\right) + \tilde{p}^T c \\ &= \max_{x\geq 0} \sum_i U(x_i) - \sum_i x_i \left(\min_{r^i\in\mathcal{H}^i} \tilde{p}^T r^i\right) + \tilde{p}^T c \\ &= g(\tilde{p}) \\ &\geq g(p^*) \end{aligned}$$

where the third equality follows from (17). Therefore, $f(\tilde{p}) = g(p^*) = g(\tilde{p})$ and $L(\tilde{R}, \tilde{x}, \tilde{p}) = L(R^*, x^*, p^*)$. Moreover, $(\tilde{R}, \tilde{x}, \tilde{p})$ is an optimal solution of the dual problem.

Step 2: Sufficiency. Assume that there is no duality gap and (R^*, x^*, p^*) is an optimal solution for both the primal problem

[4] One can exchange the order of min and max in (18) since given $\tilde{R}$, there is no duality gap in $\max_{x\geq 0} \sum_i U_i(x_i)$ s.t. $\tilde{R}x \leq c$.

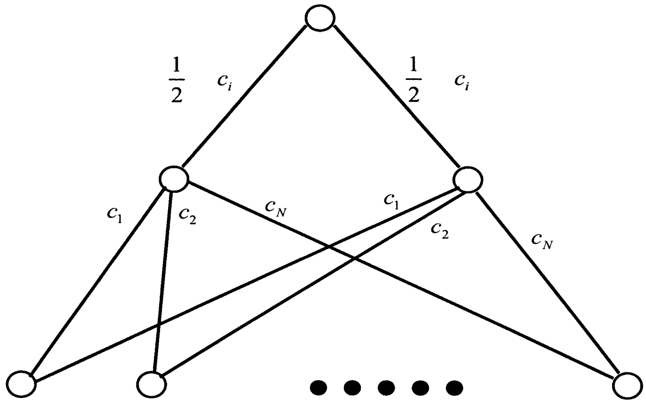


Fig. 1. Network to which integer partition problem can be reduced.

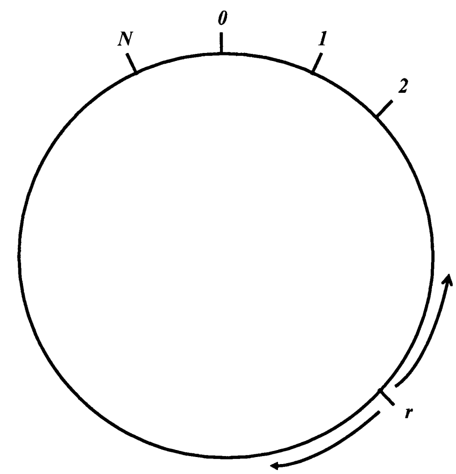


Fig. 2. A ring network.

(12) and its dual (13). We claim that it is also an equilibrium of (4)–(7) with $a = 1$ and $b = 0$, i.e., we need to show that

$$(p^*)^T r^{*i} = \min_{r^i \in \mathcal{H}^i} (p^*)^T r^i \tag{20}$$

$$(p^*, x^*) = \arg\min_{p \geq 0} \max_{x \geq 0} L(R^*, x, p)$$

$$= \arg\max_{x \geq 0} \min_{p \geq 0} L(R^*, x, p) \tag{21}$$

where r^{*i} are the ith columns of R^*. The second equality in (21) follows from the fact that there is no duality gap for the TCP-AQM problem.

Since (R^*, x^*, p^*) solves the dual problem (13), the optimal routing matrix R^* satisfies (20) by the saddle point theorem [3]. But (R^*, x^*, p^*) also solves the primal problem (12). In particular, (x^*, p^*) solves the utility maximization problem over source rates and its Lagrangian dual, with R^* as the routing matrix, i.e., (x^*, p^*) satisfies (21). □

Proof of Theorem 3: We describe a polynomial time procedure that reduces an instance of integer partition problem ([9], pp. 47) to a special case of the primal problem. Given a set of integers $c_1, \ldots, c_N$, the integer partition problem is to find a subset $A \subset \{1, \ldots, N\}$ such that

$$\sum_{i \in A} c_i = \sum_{i \notin A} c_i.$$

Given an instance of the integer partition problem, consider the network in Fig. 1, with N sources at the root, two relay nodes, and N receivers, one at each of the N leaves. The two links from the root to the relay nodes have a capacity of $\sum_i c_i/2$ each, and the two links from each relay node to receiver i have a capacity of c_i. All receivers have the same utility function that is increasing. The routing decision for each source is to decide which relay node to traverse. Clearly, maximum utility of $\sum_i U_i(c_i)$ is attained when each receiver i receives at rate c_i, from exactly one of the relay nodes, and the links from the root to the two relay nodes are both saturated. Such a routing exists if and only if there is a solution to the integer partition problem. □

C. Remark: $b > 0$ Case

The case of $b > 0$ for general network is completely open. If $a = 0$ and $b > 0$, routing $R(t) = R$, for all t, is the static minimum-cost routing with τ_l as the link costs. An equilibrium $(R, x(R), p(R))$ always exists in this case. Even though R minimizes routing cost and $(x(R), p(R))$ solves (4)–(5), it is not known if $(R, x(R), p(R))$ *jointly* solves any optimization problem.

For the case of $a > 0$ and $b > 0$, even the existence of equilibrium is unknown for general networks.

IV. Stability of TCP/IP

Theorem 1 suggests using pure prices $p(t)$ generated by TCP-AQM as link costs because in this case, an equilibrium of TCP/IP, when it exists, maximizes aggregate utility over both rates and routes. We show in this section however that the equilibrium can be unstable. Routing can be stabilized by including a strictly positive traffic-insensitive component in the link cost ($b > 0$), but at a reduced achievable utility. There thus seems to be an inevitable tradeoff between achievable utility and routing stability.

To make this precise, we analyze a special ring network with a common destination. As remarked in the last section, for a general network, we do not even know if an equilibrium exists when $b > 0$, let alone characterizing it. For the ring network, however, not only does equilibrium always exists (if we ignore the integral gap), but we can also study rigorously its stability and achievable utility, and their tradeoff under minimum-cost routing. We illustrate through a numerical example that the qualitative conclusions derived from the ring network seem to generalize to a general network.

A. Ring Network

Consider a ring network with $N + 1$ nodes, indexed by $i = 0, 1, \ldots, N$. Nodes $i \geq 1$ are sources and their common destination is node 0; see Fig. 2. For notational convenience we will also refer to node 0 as node $N + 1$. Each pair of nodes is connected by two links, one in each direction. We will refer to the two unidirectional links between node $i-1$ and i as link i; the direction should be clear from the context. The fixed delay on link i is denoted as $\tau_i > 0, i = 1, \ldots, N + 1$, in each direction. As before, the cost on link i in period t is $d_i(t) = ap_i(t) + b\tau_i$ where $p_i(t)$ is the price on link i. At time t, source i routes all its traffic in the direction, counterclockwise or clockwise, with the smaller cost. The ring network is particularly simple because the routing of the whole network can be represented by a single number r. Note that under minimum-cost routing, if node i sends in the counterclockwise direction, so must node $i-1$, and if node i sends in the clockwise direction, so must node $i+1$. Hence, we can represent routing on the network by $r \in \{0, \dots, N\}$ with the interpretation that nodes $1, \dots, r$ send in the counterclockwise direction and nodes $r+1, \dots, N$ send in the clockwise direction.

For this special case, we now show that the duality gap is trivial, that minimum-cost routing based just on prices ($b = 0$) indeed solves the primal and dual problems as Theorem 1 guarantees, but the equilibrium is unstable. Using a continuous model, we then show that routing can be stabilized if the weight b on the fixed delay is nonzero and the weight a on price is small enough. The maximum achievable utility however decreases with smaller a. There is thus an inevitable tradeoff between utility maximization and routing stability.

B. Utility and Stability of Pure Dynamic Routing

Suppose all sources i have the same utility function $U(x_i)$, and all links have the same capacity of $c = 1$ unit. We assume that U is *strictly* concave increasing and differentiable. Then at any time, only link 1, in the counterclockwise direction, and link $N+1$, in the clockwise direction, can be saturated and have strictly positive price. The utility maximization problem (12) reduces to the following simple form:

$$\max_{r \in \{0,\dots,N\}} \max_{x_i} \quad \sum_i U(x_i) \tag{22}$$

$$\text{subject to} \quad \sum_{i=1}^{r} x_i \le 1 \quad \text{and} \quad \sum_{i=r+1}^{N} x_i \le 1. \tag{23}$$

When routing is r, nodes $i = 1, \dots, r$ see price $p_1(r)$ on their paths while nodes $i = r+1, \dots, N$ see price $p_{N+1}(r)$ on their paths. Since these rates $x_i(r)$ and prices $p_i(r)$ are primal and dual optimal, they satisfy [19]

$$U'(x_i(r)) = p_1(r) \quad \text{for } i = 1, \dots, r, \tag{24}$$

$$U'(x_i(r)) = p_{N+1}(r) \quad \text{for } i = r+1, \dots, N. \tag{25}$$

This implies that $x_1(r) = \cdots = x_r(r)$ and $x_{r+1}(r) = \cdots = x_N(r)$.

It is easy to see that the optimal routing $r^* \neq 0$ or N. Hence both constraints are active at optimality, implying that [from (23)]

$$x_1(r) = \cdots = x_r(r) = \frac{1}{r} \tag{26}$$

$$x_{r+1}(r) = \cdots = x_N(r) = \frac{1}{N-r}. \tag{27}$$

The problem (22)–(23) thus becomes

$$\max_{r \in \{1,\dots,N-1\}} rU\left(\frac{1}{r}\right) + (N-r)U\left(\frac{1}{N-r}\right).$$

Dividing the objective function by N and using the strict concavity of U, we have

$$\frac{r}{N}U\left(\frac{1}{r}\right) + \frac{N-r}{N}U\left(\frac{1}{N-r}\right) \le U\left(\frac{2}{N}\right)$$

with equality if and only if $r = N/2$. This implies that the optimal routing is

$$r^* := \lfloor N/2 \rfloor \tag{28}$$

and the maximum utility is

$$V^* := \left\lfloor \frac{N}{2} \right\rfloor U\left(\frac{1}{\lfloor N/2 \rfloor}\right) + \left\lceil \frac{N}{2} \right\rceil U\left(\frac{1}{\lceil N/2 \rceil}\right) \tag{29}$$

where $\lfloor y \rfloor$ is the largest integer less or equal to y and $\lceil y \rceil$ is the smallest integer greater or equal to y.

It can be shown that there is no duality gap for the ring network considered here when N is even, by verifying that routing r^* in (28), rates $x_i(r^*)$ in (27), and prices $p_1(r^*), p_{N+1}(r^*)$ in (24)–(25) are indeed primal-dual optimal.[5] When N is odd, there is generally a duality gap due to integral constraint on r; see Appendix VII.A for a proof. This duality gap disappears in the convexified problem when routing is allowed to take real value in $[0, N]$, a model we consider in the next subsection. This suggests that TCP together with minimum-cost routing based on prices can potentially maximize utility for this ring network. We next show, however, that minimum-cost routing based only on prices is unstable.

Given routing r, we can combine (24)–(25) and (27) to obtain the prices $p_1(r)$ and $p_{N+1}(r)$ on links 1 and $N+1$:

$$p_1(r) = U'\left(\frac{1}{r}\right) \quad \text{and} \quad p_{N+1}(r) = U'\left(\frac{1}{N-r}\right). \tag{30}$$

The path cost for node i in the counterclockwise direction is

$$D^-(i;r) = \sum_{j=1}^{i} b\tau_j + ap_1(r) = b\sum_{j=1}^{i} \tau_j + aU'\left(\frac{1}{r}\right) \tag{31}$$

and the path cost in the clockwise direction is

$$D^+(i;r) = \sum_{j=i+1}^{N+1} b\tau_j + ap_{N+1}(r) = b\sum_{j=i+1}^{N+1} \tau_j + aU'\left(\frac{1}{N-r}\right). \tag{32}$$

In the next period, each node i will choose counterclockwise or clockwise direction according as $D^-(i;r)$ or $D^+(i;r)$ is smaller. Define $f(r)$ as

$$f(r) := \max\{i \mid D^-(i;r) \le D^+(i;r)\}. \tag{33}$$

Then the resulting routing satisfies the recursive relation

$$r(t+1) = \begin{cases} 0, & \text{if } D^-(1;r(t)) > D^+(1;r(t)) \\ N, & \text{if } D^-(N;r(t)) < D^+(N;r(t)) \\ f(r(t)), & \text{otherwise.} \end{cases}$$

[5]This also follows from Theorem 1 and the fact that $r = N/2$ is by symmetry the equilibrium routing when N is even.

Theorem 4: If $b = 0$ and $a > 0$, then starting from any routing $r(0)$, except the equilibrium $N/2$ when N is even, the subsequent routing oscillates between 0 and N.

Proof: For any $r(0) \in \{0, \ldots, N\}$, we have

$$\begin{aligned} & D^-(1; r(0)) - D^+(1; r(0)) \\ & = D^-(N; r(0)) - D^+(N; r(0)) \\ & = a \left(U' \left(\frac{1}{r(0)} \right) - U' \left(\frac{1}{N - r(0)} \right) \right). \end{aligned}$$

If N is even, then $N/2$ is the unique equilibrium routing that solves $D^-(i; N/2) = D^+(i; N/2)$. Suppose $r(0) \neq N/2$. If $r(0) > N/2$, then $1/r(0) < 2/N < 1/(N - r(0))$. Since U' is strictly decreasing, $U'(1/r(0)) > U'(1/(N - r(0))$ and hence $D^-(1; r(0)) > D^+(1; r(0))$ and $r(1) = 0$. Similarly, if $r(0) < N/2$, then $D^-(N; r(0)) < D^+(N; r(0))$ and $r(1) = N$. Hence r oscillates between 0 and N henceforth. □

Even though purely dynamic routing based on prices ($b = 0$) maximizes utility, Theorem 4 says that it is unstable. We will henceforth, without loss of generality, set $b = 1$ and consider the effect of a on utility maximization and stability.

C. Maximum Utility of Minimum-Cost Routing

As mentioned above, the duality gap for the ring network we consider is of a trivial kind that disappears when integer constraint on routing is relaxed. For the rest of this section, we consider a continuous model where every point on the ring is a source. A point on the ring is labeled by $s \in [0, 1]$ and the common destination is the point 0 (or equivalently 1). The utility maximization problem becomes

$$\max_{r \in [0,1]} \max_{x(\cdot)} \quad \int_0^1 U(x(u))\, du \tag{34}$$

$$\text{subject to} \quad \int_0^r x(u)\, du \le 1 \tag{35}$$

$$\int_r^1 x(u)\, du \le 1. \tag{36}$$

As in the discrete case, both constraints are active at optimality, and hence the problem reduces to

$$\max_{r \in (0,1)} rU\left(\frac{1}{r}\right) + (1 - r)U\left(\frac{1}{1 - r}\right)$$

which, by concavity, yields the optimal routing r^* and maximum utility V^*

$$r^* = \frac{1}{2} \quad \text{and} \quad V^* = U(2). \tag{37}$$

To see that there is no duality gap, note that the problem (34)–(36) is equivalent to:

$$\begin{aligned} \max_{r \in [0,1]} \max_{x^-, x^+ \ge 0} \quad & rU(x^-) + (1 - r)U(x^+) \\ \text{subject to} \quad & rx^- \le 1, \quad (1 - r)x^+ \le 1. \end{aligned}$$

Define the Lagrangian as

$$\begin{aligned} & L(r, x^-, x^+, p^-, p^+) \\ & \quad = rU(x^-) + (1 - r)U(x^+) + p^-(1 - rx^-) \\ & \qquad + p^+(1 - (1 - r)x^+). \end{aligned}$$

It is easy to verify that

$$r^* = \frac{1}{2}, \quad x^{-*} = x^{+*} = 2, \quad p^{-*} = p^{+*} = U'(2). \tag{38}$$

are primal-dual optimal and there is no duality gap; see Appendix VII-B.

We now look at the maximum utility achievable by the equilibrium of minimum-cost routing.

Let the delay from s to the destination in the counterclockwise direction be

$$T(s) := \int_0^s \tau(u)\, du$$

and the delay in the clockwise direction be

$$T(1) - T(s) = \int_s^1 \tau(u)\, du$$

where $\tau(u), u \in [0, 1]$, is given. Here, $\tau(u)$ corresponds to link cost in the discrete model. Given routing $r \in [0, 1]$, the price in the counterclockwise direction is $U'(1/r)$ and the price in the clockwise direction is $U'(1/(1 - r))$. Then the cost of source s in the counterclockwise direction is

$$D^-(s; r) = T(s) + aU'\left(\frac{1}{r}\right) \tag{39}$$

and the cost in the clockwise direction is

$$D^+(s; r) = T(1) - T(s) + aU'\left(\frac{1}{1 - r}\right). \tag{40}$$

A routing r is in equilibrium if the costs of source r in both directions are the same.

Definition 5: A routing r is called an *equilibrium routing* if $D^-(r; r) = D^+(r; r)$. It is denoted by r_a or $r(a)$.

By definition, r_a is the solution of

$$\begin{aligned} g(r) &:= 2T(r) - T(1) + a\left(U'\left(\frac{1}{r}\right) - U'\left(\frac{1}{1 - r}\right)\right) \\ &= 0. \end{aligned} \tag{41}$$

Since $g(0) < 0, g(1) > 0$, and $g'(r) > 0$, the equilibrium r_a is in (0, 1) and is unique.

Given a routing r, its utility is

$$V(r) := rU\left(\frac{1}{r}\right) + (1 - r)U\left(\frac{1}{1 - r}\right).$$

The maximum utility achieved by minimum-cost routing, with parameter a, is then $V(r_a) \le V(r^*) = V^*$. The next result implies that r_a varies between r_0 and r^* and converges monotonically to r^* as $a \to \infty$. As a result, the loss $V^* - V(r_a) \ge 0$ in utility also approaches 0 as $a \to \infty$. Denote the interval in which $1/r_a$ and $1/(1-r_a)$ vary as $I := [2, 1/\min\{r_0, 1-r_0\}]$.

Theorem 6: Suppose U'' exists and is bounded on I. For all $a \geq 0, |r_a - r^*|$ is a strictly decreasing function of a. Moreover, as $a \to \infty, |r_a - r^*|$ and $V^* - V(r_a)$ approach 0.

Proof: The (41) defines the equilibrium routing $r(a) := r_a$ as an implicit function of a. By the implicit function theorem, $r'(a)$ satisfies

$$\frac{1}{r'(a)}\left[U'\left(\frac{1}{1-r_a}\right) - U'\left(\frac{1}{r_a}\right)\right] = 2\tau(r_a) - \frac{a}{r_a^2}U''\left(\frac{1}{r_a}\right) - \frac{a}{(1-r_a)^2}U''\left(\frac{1}{1-r_a}\right).$$

The right-hand side is positive since U is strictly concave. Hence $r'(a)$ has the same sign as the term in the square bracket, i.e., since U' is decreasing,

$$r'(a) = \begin{cases} >0, & \text{if } r_a < r^* \\ <0, & \text{if } r_a > r^* \\ =0, & \text{if } r_a = r^*. \end{cases} \tag{42}$$

This implies that $|r_a - r^*|$ is a strictly decreasing function of a; see Fig. 3.

Hence $|r_a - r^*|$ converges to a limit as $a \to \infty$. Since U'' is bounded on the closed interval I, so is U'. Hence, from (41), we must have

$$U'(1/r_a) - U'(1/(1-r_a)) \to 0$$

or

$$U'(1/\lim_{a\to\infty} r_a) = U'(1/(1 - \lim_{a\to\infty} r_a)).$$

Since U' is strictly decreasing, this implies that $\lim_{a\to\infty} r_a = 1 - \lim_{a\to\infty} r_a = r^*$.

To show that $V^* - V(r_a) \geq 0$ also converges to 0, note that $V'(r^*) = 0$ and hence we have, by Taylor expansion,

$$V(r_a) - V^* = \frac{1}{2}V''(u)(r_a - r^*)^2$$

for some u between r_a and r^*. Here

$$V''(u) = \frac{1}{u^3}U''\left(\frac{1}{u}\right) + \frac{1}{(1-u)^3}U''\left(\frac{1}{1-u}\right) \geq -\frac{2\mu}{(\min\{r_0, 1-r_0\})^3}$$

where μ is the upper bound of U'' on I. Hence

$$0 \leq V^* - V(r_a) \leq \frac{\mu(r_a - r^*)^2}{(\min\{r_0, 1-r_0\})^3}.$$

Since $|r_a - r^*| \to 0$, the proof is complete. □

The shape of $r'(a)$ in (42) implies that, if $r(0) > r^*$ then $r(a) \geq r^*$ for all a but $r(a)$ decreases to r^* as $a \to \infty$, and if $r(0) < r^*$ then $r(a) \leq r^*$ for all a but $r(a)$ increases to r^* monotonically, as illustrated in Fig. 3. This is a consequence of the continuity of $r(a)$.

D. Stability of Minimum-Cost Routing

We now turn to the stability of r_a. For simplicity, we will take $U(x) = \log x$, the utility function of TCP Vegas [21]. With log utility function, $V'(r_a) = \log(1-r)/r$ and hence Theorem 6 can be strengthened to show that $V^* - V(r_a)$ is a strictly decreasing function of a, and hence converges monotonically to 0 as $a \to \infty$.

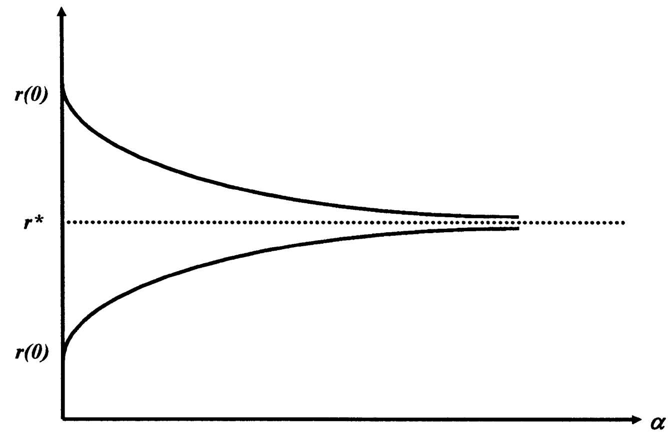


Fig. 3. $r(a)$.

Given r, let $f(r)$ denote the solution of

$$D^-(s;r) = D^+(s;r).$$

It is in the range [0, 1] if and only if $0 \leq T(s) \leq T(1)$, or if and only if

$$r^* - \frac{T(1)}{2a} \leq r \leq r^* + \frac{T(1)}{2a}.$$

We will assume that $\min_{u\in[0,1]} \tau(u) > 0$. Then T^{-1} exists and

$$f(r) = T^{-1}\left(\frac{1}{2}(T(1)+a) - ar\right). \tag{43}$$

The routing iteration is

$$r(t+1) = [f(r(t))]_0^1 \tag{44}$$

where $[r]_0^1 = \max\{0, \min\{1, r\}\}$.

Definition 7: The equilibrium routing r_a is *(globally) stable* if starting from any routing $r(0), r(t)$ defined by (43)–(44) converges to r_a as $t \to \infty$.

Example 2: Uniform τ: Suppose delay is uniform on the ring, $\tau(u) = \tau$ for all $u \in [0,1]$, so that $T(r) = r\tau$. From (41), the equilibrium routing is

$$r_a = \frac{1}{2} = r^*, \quad \forall a \geq 0$$

coinciding with the utility-maximizing routing r^*.

Suppose $a < \tau$. Then the routing iteration becomes

$$r(t+1) = \frac{1}{2\tau}(\tau + a) - \frac{a}{\tau}r(t) = f(r(t)).$$

Since $|f(s) - f(r)| = (a/\tau)|s - r| < |s - r|, f(r)$ is a contraction mapping and hence r_a is globally stable for all $0 \leq a < \tau$. □

Hence for the uniform delay case, adding a static component to link cost stabilizes routing provided the weight on prices is smaller than link delay. Moreover, the static component does not lead to any loss in utility ($r_a = r^*$). The stability condition generalizes to the general delay case. The following theorem says that if a is smaller than the minimum 'link delay', then r_a is globally stable; if a is bigger than the maximum 'link delay', then it is globally unstable (diverge from any initial routing except r_a); otherwise, it may converge or diverge depending on initial routing.

Theorem 8:

1) If $a < \min_{u\in[0,1]} \tau(u)$ then r_a is globally stable.
2) Suppose $a \geq T(1)$. Then there exists $\underline{r} < r_a < \bar{r}$ such that
a) If $r(0) = \underline{r}$ or $r(0) = \bar{r}$ then subsequent routings oscillate between $\bar{r}$ and $\underline{r}$.
b) If $r(0) < \underline{r}$ or $r(0) > \bar{r}$ then subsequent routings after a finite number of iterations oscillate between 0 and 1.
c) If $\underline{r} < r(0) < \bar{r}$ then $r(t)$ converges to r_a provided $a < \min_{u\in(\underline{r},\bar{r})} \tau(u)$.
3) If $a > \max_{u\in[0,1]} \tau(u)$ then starting from any initial routing $r(0) \neq r_a$, subsequent routings after a finite number of iterations oscillate between 0 and 1.

Proof:

1. We show that the routing iteration (44) is a contraction mapping if $a < \min_{u\in[0,1]} \tau(u)$. Now

$$\begin{aligned}
&\left|[f(s)]_0^1 - [f(r)]_0^1\right| \\
&\quad \leq |f(s) - f(r)| \\
&\quad = \left|T^{-1}\left(\frac{T(1)+a-2as}{2}\right) - T^{-1}\left(\frac{T(1)+a-2ar}{2}\right)\right| \\
&\quad = \left|\frac{1}{T'(u)}\,(as - ar)\right| \\
&\quad \leq \frac{a}{\min_{u\in[0,1]} \tau(u)}|s-r|
\end{aligned}$$

for some u between r and s, by the mean value theorem. Hence $h(r)$ is a contraction mapping and starting from any $r(0) \in [0,1], r(t)$ converges exponentially to r_a.

2. Define

$$h(r) = \frac{1}{2}(T(1)+a) - ar.$$

Then the routing iteration can be written as

$$T(r(t+1)) = [h(r(t))]_0^1. \tag{45}$$

Define the following sequences:

$$\begin{aligned}
a_0 &= 0, \quad b_0 = T(0) \\
a_{n+1} &= h^{-1}(b_n), \quad b_{n+1} = T(a_{n+1}).
\end{aligned}$$

Note that $(a_n, n \geq 0)$ is a routing sequence going backward in time. The following lemma is proved in the Appendix, following [17].

Lemma 9: Let $T_a = T(r_a) = h(r_a)$. Then

$$\begin{aligned}
0 = a_0 < a_2 < \cdots < r_a < \cdots < a_3 \ < \ a_1 < 1 \\
T(0) = b_0 < b_2 < \cdots < T_a < \cdots < b_3 \ < \ b_1 < T(1)
\end{aligned}$$

Since the sequences are monotone, the lemma implies that there are $\underline{r}$ and $\bar{r}$ with $0 < \underline{r} < r_a < \bar{r} < 1$ such that

$$\lim_{n\to\infty} a_{2n} = \underline{r} \quad \text{and} \quad \lim_{n\to\infty} a_{2n+1} = \bar{r}.$$

By continuity of T and h, we have

$$T(\underline{r}) = h(\bar{r}) \quad \text{and} \quad T(\bar{r}) = h(\underline{r}).$$

This implies that starting from $r(0) = \underline{r}$ or $r(0) = \bar{r}$, the subsequent routings oscillate between $\underline{r}$ and $\bar{r}$.

To show the second claim, suppose $r(0) < \underline{r}$. Specifically, suppose $a_{2n-2} < r(0) < a_{2n}$ for some n. If $h(r(0)) > T(1)$ (possible since $a \geq T(1)$), then $r(1) = 1$ and subsequent routings oscillate between 0 and 1. Otherwise, from (45), $r(0) = h^{-1}(T(r(1)))$, and hence $a_{2n-2} < h^{-1}(T(r(1))) < a_{2n}$. Since h is strictly decreasing, we have $b_{2n-1} < T(r(1)) < b_{2n-3}$ by definition of b_n. Hence, since T is strictly increasing, $a_{2n-1} < r(1) < a_{2n-3}$. The same argument then shows that $a_{2n-4} < r(2) < a_{2n-2}$. Hence we have shown that $r(0) < a_{2n}$ implies $r(2) < a_{2n-2}$. This proves the second claim.

The proof of the third claim follows the same argument of part 1.

3. By the mean value theorem, we have, for all α, α' in [0, 1],

$$|h^{-1}(T(\alpha)) - h^{-1}(T(\alpha'))| = \frac{T'(u)}{a}|\alpha - \alpha'|$$

for some u between α and α'. Hence the iteration map

$$a_{n+1} = h^{-1}(T(a_n))$$

is a contraction provided $a > \max_{u\in[0,1]} \tau(u)$. This implies that the sequence $(a_n, n \geq 0)$ converges and, since r_a is the unique fixed point of $h^{-1}(T(\cdot)), \underline{r} = \bar{r} = r_a$. The assertion then follows from part 2(b). □

E. General Network

It seems difficult to derive an analytical bound on a to guarantee routing stability or to compute optimal routing for general networks. In this section, we present numerical results to illustrate that the intuition from the simple ring network analyzed in the previous subsections extends to general topology.

We generate a random network based on Waxman's algorithm [29]. The nodes are uniformly distributed in a two dimensional plane. The probability that a pair of nodes u, v are connected is given by

$$\text{Prob}(u,v) = \alpha \exp\left(\frac{d(u,v)}{\beta L}\right)$$

where the maximum probability $\alpha > 0$ controls connectivity, $\beta \leq 1$ controls the length of the edges with a larger β favoring longer edges, $d(u,v)$ is the Euclidean distance between nodes u and v, and L is the maximum distance between any two nodes. In our example, we set the number of nodes $N = 30, \alpha = 0.8, \beta = 0.3$, which generated 90 bidirectional links; see Fig. 4. The fixed delay τ_l of each link l is randomly chosen according to a uniform distribution over [100, 400] ms. The link capacities are randomly chosen from the interval [1000, 4000] packets/sec, also with uniform distribution. There are 60 flows on the network with randomly chosen source and destination nodes. Routing on this network is computed using Bellman-Ford minimum-cost algorithm, with link cost $d_l(t) = \tau_l + ap_l(t)$

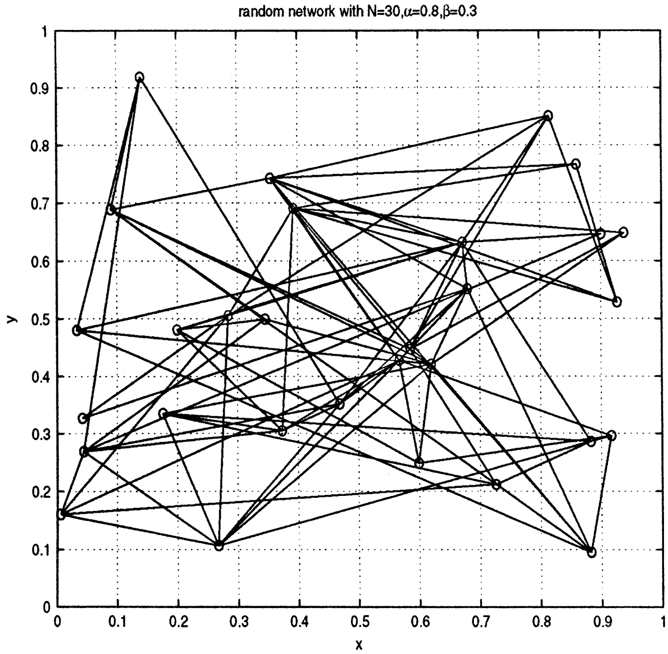


Fig. 4. A random network.

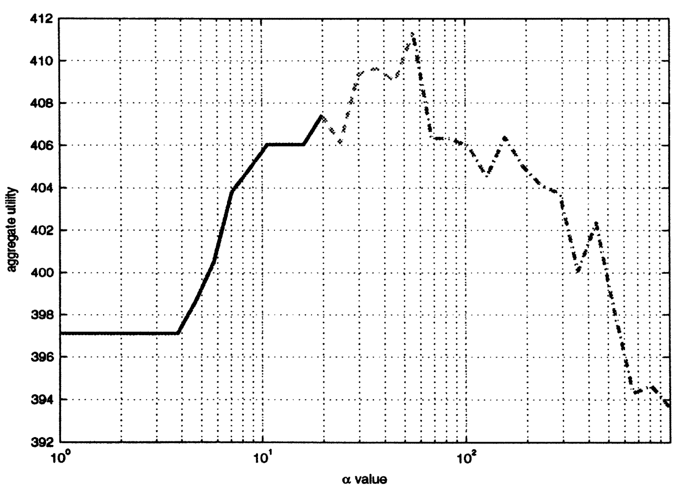


Fig. 5. Aggregate utility as a function of a for the random network in Fig. 4.

in each update period t, on a slower timescale than congestion control. In each routing period t, we first solve the link prices based on the current routing, using the gradient projection algorithm of [19]. We iterate the source algorithm to update rates and the link algorithm to update prices, until they converge. The link prices are then used to compute the minimum-cost for the next period.

We measure the performance of the scheme at different a by the sum of all source's utilities. If the routing is stable (at small a), the aggregate utility is computed using the equilibrium routing. Otherwise, the routing oscillates and the time-averaged aggregate utility is used. The result is shown in Fig. 5.

As expected, when a is small, routing is stable and the aggregate utility increases with a, as in the ring network analyzed in Section IV-C (Theorem 6). When $a < 4$, the static delay τ_l dominates the link cost and the routes computed with $d_l(t)$ remain the same as with static routing ($a = 0$), and hence the aggregate utility is independent of a. Routing becomes unstable at around $a = 10$. Even though the time-averaged utility continues to rise after routing instability sets in, eventually it peaks and drops to a level less than the utility achievable by the necessarily stable static routing.

V. Resource Provisioning

Results in the previous sections show that even though an equilibrium of TCP/IP, when it exists, maximizes utility under pure dynamic routing, it can be unstable and hence not attainable by the TCP/IP system. In this section, we show that if the link capacities are optimally provisioned, however, pure *static* routing is enough to maximize utility. Moreover, it is optimal even within the class of multipath routing: again, there is no penalty in not splitting traffic across multiple paths.

Suppose it costs $\alpha_l > 0$ amount to provision a unit of capacity at link l and let $\alpha = (\alpha_l, \text{for all } l)$ be the vector of unit costs. For instance, a longer link may have a larger α_l. The total capacity cost over the entire network is $\alpha^T c$. Suppose the budget for provisioning the capacity is B. Consider the problem of optimally selecting capacities, routing, and source rates to maximize utility:

$$\max_{c\geq 0} \max_{R\in\mathcal{R}_m} \max_{x\geq 0} \quad \sum_i U_i(x_i) \tag{46}$$

$$\text{subject to} \quad Rx \leq c \tag{47}$$

$$\alpha^T c \leq B. \tag{48}$$

where U_i are concave increasing utility functions. Note that R ranges in $\mathcal{R}_m$, and hence multipath routing is allowed and the problem has no duality gap. This problem may arise when optical lightpaths can be dynamically reconfigured at connection timescale.

Theorem 10: Suppose $U_i'(x_i) > 0$ for all i and $x_i \geq 0$. At optimality:

1) there is an optimal solution (c^*, R^*, x^*) where $R^* \in \mathcal{R}_s$ is a single-path routing.
2) moreover, R^* is pure static routing using α_l as link costs.
3) $R^* x^* = c^*$, i.e., there is no slack capacity.
4) $\alpha^T c = B$, i.e., there is no slack in budget.
5) link prices generated by TCP-AQM are proportional to the provisioning costs, $p^* = \gamma^* \alpha$ for some $\gamma^* > 0$.

Proof: It is easy to show the existence of an equilibrium. Define the Lagrangian of (46)–(48) as

$$L(c, R, x, p, \gamma) = \sum_i U_i(x_i) - p^T(Rx - c) - \gamma(\alpha^T c - B).$$

At optimality, the KKT condition holds: there exist $p^* \geq 0$ and $\gamma^* \geq 0$ such that

$$U_i^{'}(x_i^*) = \sum_l R_{li}^* p_l^*$$

$$p^* = \gamma^* \alpha \tag{49}$$

$$(p^*)^T (R^* x^* - c^*) = 0 \tag{50}$$

$$\gamma^* (\alpha^T c^* - B) = 0. \tag{51}$$

From (49), we obtain the last claim in the theorem. Moreover, (49) and $U_i'(x_i^*) > 0$ imply that $\gamma^* > 0$ and $p_l^* > 0$ for *all* l, since $\alpha > 0$. Hence, from (50), (51) and primal feasibility, equality holds in (47) and (48), proving the third and fourth claims.

To prove the first two claims, express the routing matrix R as $R = HW$ where $W \in \mathcal{W}_m$. Using the equalities in (47) and (48) to eliminate c, we can transform the utility maximization problem (46)–(48) into

$$\begin{aligned} \max_{W \in \mathcal{W}_m} \max_{x \geq 0} \quad & \sum_i U_i(x_i) \\ \text{subject to} \quad & \sum_i (\alpha^T H^i w^i) x_i = B \end{aligned}$$

where W =diag(w^i). Since U_i is nondecreasing and both the objective and the constraints above are separable in i, in order to maximize utility, w^i should be chosen to be a solution of

$$\begin{aligned} \min_{w^i} \quad & \alpha^T H^i w^i \\ \text{subject to} \quad & \mathbf{1}^T w^i = 1, \quad 0 \leq w^i_j \leq 1. \end{aligned}$$

Since this is a linear program, there exists an optimal point on the boundary. Hence there is an optimal $W^* \in \mathcal{W}_s$, i.e., minimum-cost single-path routing using α_l as link costs is optimal. □

VI. Conclusion

Given a routing, TCP-AQM can be interpreted as a distributed primal-dual algorithm over the Internet to maximize aggregate utility over source rates. In this paper, we study whether TCP-AQM together with IP (modeled by minimum-cost routing) can maximize utility over both source rates and routing, on a slower timescale. We show that we can indeed interpret TCP/IP as *attempting* to maximize utility in the sense that its equilibrium, if exists, solves the utility maximization problem and its dual, provided congestion prices generated by TCP-AQM are used as link costs. TCP/IP equilibrium exists if and only if there is no penalty in not splitting traffic across multiple paths. Even if equilibrium exists, however, TCP/IP with pure dynamic routing can be unstable. Specializing to a special ring network, we show that routing is indeed unstable when link costs are congestion prices, it can be stabilized by adding a static component to the definition of link cost, but the static component reduces the achievable utility. There thus seems to be an inevitable tradeoff between routing stability and utility maximization, for any given set of link capacities. We show, however, that if link capacities are optimally provisioned, then pure static (and hence stable) routing is sufficient to maximize utility even for general networks and link costs are proportional to the provisioning costs. Moreover single-path routing can achieve the same utility as multipath routing. Hence, one can regard the layering of TCP and IP as a decomposition of the utility maximization problem over source rates and routes into a distributed and decentralized algorithm, carried out on different timescales, at least when network capacities are well provisioned.

The duality model of TCP-AQM has been useful in understanding the equilibrium properties, including throughput, packet loss, delay, and fairness, of large-scale networks under TCP-AQM control. This paper is a first, and preliminary, attempt to apply the same methodology to understand the cross-layer interaction of TCP-AQM, minimum-cost routing and resources allocation. Our model is simplistic—it ignores finite duration flows and randomness in real networks, and reduces the rich behavior of IP to minimum-cost routing. Even within this highly abstract model, many questions remain open. First, even though numerical examples suggest that the tradeoff between routing stability and utility maximization is present in a more general network than the special ring network we studied, we have not been able to find an analytical proof. One of the major difficulties is that, in a general network, minimum-cost routing cannot be as conveniently represented as in the ring network. Second, when static component is included in link cost $b > 0$, it is not known if TCP/IP has an equilibrium, whether the equilibrium jointly solves a certain optimization problem, and under what condition it is stable. Third, it would be interesting to estimate the duality gap in the single-path problem. Even though this problem is not directly related to the TCP/IP iteration when the duality gap is nonzero, the gap measures the penalty of not splitting traffic among multiple paths.

Appendix

A. Duality Gap When N Is Odd

We prove that there is generally a duality gap between the primal problem (22)–(23) and its dual when N is odd.

It is easy to see that the primal optimal routing is

$$r^* = \frac{N-1}{2} \quad \text{or} \quad \frac{N+1}{2}.$$

Suppose without loss of generality that $r^* = (N-1)/2$ (the other case is similar). Then, the source rates are

$$x_1 = \cdots = x_{r^*} = \frac{2}{N-1}$$

and

$$x_{r^*+1} = \cdots = x_N = \frac{2}{N+1}$$

yielding a primal objective value of

$$\begin{aligned} &\frac{N-1}{2} U\left(\frac{2}{N-1}\right) + \frac{N+1}{2} U\left(\frac{2}{N+1}\right) \\ &\quad = N\left\{\left(\frac{1}{2} - \frac{1}{N}\right) U\left(\frac{2}{N-1}\right) + \left(\frac{1}{2} + \frac{1}{N}\right) U\left(\frac{2}{N+1}\right)\right\} \\ &\quad < NU\left(\frac{2}{N}\right) \end{aligned}$$

where the last inequality follows from the strict concavity of U. We now show that the right-hand side is the optimal dual objective value, and hence there is a duality gap.

The dual problem of (22)–(23) is (e.g., [19])

$$\min_{p_1, p_{N+1} \geq 0} \left(\sum_{i=1}^{N} \max_{x_i} \ \phi(x_i, p_1, p_{N+1}) + (p_1 + p_{N+1}) \right)$$

where

$$\phi(x_i, p_1, p_{N+1}) = U(x_i) - x_i \ \min\{p_1, p_{N+1}\}.$$

First, note that the minimizing (p_1, p_{N+1}) must satisfy $p_1 = p_{N+1}$, for otherwise, if (say) $p_1 < p_{N+1}$, then the dual objective value is

$$\sum_{i=1}^{N} \max_{x_i} \left(U(x_i) - x_i p_1\right) + (p_1 + p_{N+1})$$

and can be reduced by decreasing p_{N+1} to p_1. Hence the dual problem is equivalent to

$$\min_{p \geq 0} \quad \sum_{i=1}^{N} \max_{x_i} (U(x_i) - x_i p) + 2p. \tag{52}$$

Let p^* denote the minimizer and $x_i^* = x_i(p^*) = x(p^*) =: x^*$ denote the corresponding maximizers (they are equal for all i by symmetry). Then we have

$$U'(x^*) = p^*. \tag{53}$$

Differentiating the objective function in (52) with respect to p and setting it to zero, we have

$$0 = N(U'(x^*)x'(p^*) - p^* x'(p^*) - x^*) + 2. \tag{54}$$

Using (53), we have

$$x^* = \frac{2}{N}$$

and hence the minimum dual objective value is

$$N(\max_{x^*} U(x^*) - x^* p^*) + 2p^* = NU\left(\frac{2}{N}\right)$$

as desired. □

B. Primal-Dual Optimality

We prove that the solution given by (38) is primal-dual optimal using the saddle-point theorem (e.g., ([3], pp. 427)). Clearly, (r^*, x^{-*}, x^{+*}) is primal feasible and (p^{-*}, p^{+*}) is dual feasible. We now show that $(r^*, x^{-*}, x^{+*}, p^{-*}, p^{+*})$ is a saddle point, i.e., for all (r, x^-, x^+, p^-, p^+),

$$\begin{aligned} L(r, x^-, x^+, p^{-*}, p^{+*}) &\leq L(r^*, x^{-*}, x^{+*}, p^{-*}, p^{+*}) \\ &\leq L(r^*, x^{-*}, x^{+*}, p^-, p^+). \end{aligned} \tag{55}$$

For the right inequality, substitute (r^*, x^{-*}, x^{+*}) from (38) into $L(r^*, x^{-*}, x^{+*}, p^-, p^+)$ to get, for all (p^-, p^+)

$$L(r^*, x^{-*}, x^{+*}, p^-, p^+) = U(2).$$

But $U(2) = L(r^*, x^{-*}, x^{+*}, p^{-*}, p^{+*})$, establishing the right inequality.

For the left inequality, denoting $p^* := p^{-*} = p^{+*}$, from (38) we have

$$\begin{aligned} &L(r, x^-, x^+, p^{-*}, p^{+*}) \\ &\quad = rU(x^-) + (1-r)U(x^+) - (rx^- + (1-r)x^+)p^* \\ &\qquad + 2p^* \\ &\quad \leq U(y) - yp^* + 2p^* \quad \text{(concavity of } U) \end{aligned} \tag{56}$$

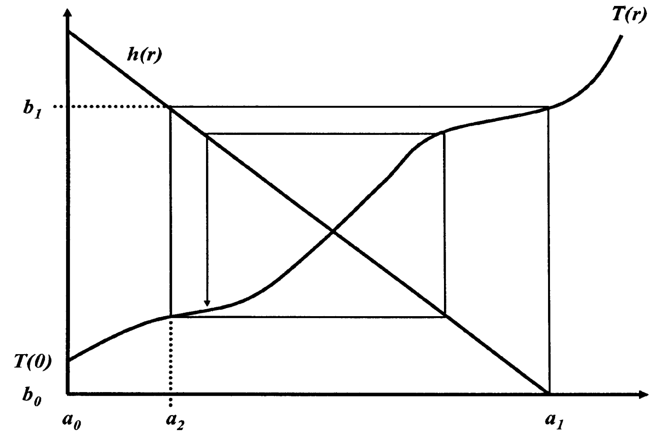


Fig. 6. Lemma 9.

with $y := rx^- + (1-r)x^+$, where equality holds if and only if $x^- = x^+$ since U is *strictly* concave. Notice that the right-hand side is maximized over y if and only if y satisfies

$$U'(y) = p^*.$$

This implies that $y = x^{-*} = x^{+*} = 2$ since U' is strictly monotonic. Substitute $y = 2$ into (56) yields, for all (r, x^-, x^+),

$$L(r, x^-, x^+, p^{-*}, p^{+*}) \leq U(2)$$

as desired, since $U(2) = L(r^*, x^{-*}, x^{+*}, p^{-*}, p^{+*})$. □

C. Proof of Lemma 9

We will prove the lemma by induction. Note that $b_0 < T_a$ implies that $a_1 = h^{-1}(b_0) > h^{-1}(T_a) = r_a$. Since $a \geq T(1)$ and $h(1) < 0, a_1 = h^{-1}(b_0) < 1$ (see Fig. 6). Hence

$$0 = a_0 < r_a < a_1 < 1.$$

This implies that $b_1 = T(a_1)$ satisfies

$$T(0) = b_0 < T_a < b_1 < T(1).$$

Since $b_1 < T(1) < h(0), a_2 = h^{-1}(b_1) > h^{-1}(h(0)) = 0$, we have

$$0 = a_0 < a_2 < r_a < a_1 < 1.$$

Let the induction hypothesis be

$$\begin{aligned} a_0 < \cdots < a_{2n} < r_a < a_{2n-1} < \cdots < a_1 \\ b_0 < \cdots < b_{2n-2} < T_a < b_{2n-1} < \cdots < b_1. \end{aligned}$$

Then $b_{2n} = T(a_{2n}) > T(a_{2n-2}) = b_{2n-2}$ and that $b_{2n} = T(a_{2n}) < T(r_a) = T_a$. Hence,

$$b_{2n-2} < b_{2n} < T_a.$$

This implies that $r_a < a_{2n+1} < a_{2n-1}$, which in turn implies that $T_a < b_{2n+1} < b_{2n-1}$. This completes the induction. □

ACKNOWLEDGMENT

The authors thank C. Umans of Caltech for the interpretation of Theorem 3. They also thank the anonymous reviewers, whose comments motivated most of the results in Section III, and for supplying several references in the theoretical computer science literature.

REFERENCES

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[2] B. Awerbuch, Y. Azar, and S. Plotkin, "Throughput competitive online routing," in *Proc. 34th IEEE Symp. Foundations of Computer Science*, 1993, pp. 32–40.

[3] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[4] D. P. Bertsekas, "Dynamic behavior of shortest path routing algorithms for communication networks," *IEEE Trans. Automat. Contr.*, vol. 27, no. 1, pp. 60–74, Feb. 1982.

[5] ——, *Linear Network Optimization: Algorithms and Codes*. Cambridge, MA: MIT Press, 1991.

[6] S. Cosares and I. Saniee, "An optimization problem related to balancing loads on SONET rings," *Telecommun. Syst.*, vol. 3, pp. 165–181, 1994.

[7] S. Deb and R. Srikant, "Congestion control for fair resource allocation in networks with multicast flows," in *Proc. IEEE Conf. Decision and Control*, Dec. 2001, pp. 1911–1916.

[8] R. G. Gallager and S. J. Golestani, "Flow control and routing algorithms for data networks," in *Proc. 5th Int Conf. Comp. Comm.*, 1980, pp. 779–784.

[9] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA: Freeman, 1979.

[10] N. Garg and J. Konemann, "Faster and simpler algorithms for multicommodity flow and other fractional packing problems," in *Proc. 39th Annu. Symp. Foundations of Computer Science*, 1998, pp. 300–309.

[11] C. Hedrick. (1988, Jun.) Routing Information Protocol. IETF RFC 1058. [Online]. Available: http://www.faqs.org/rfcs/rfc1058.html

[12] K. Kar, S. Sarkar, and L. Tassiulas, "Optimization based rate control for multipath sessions," in *Proc. 7th Int. Teletraffic Congress (ITC)*, Dec. 2001.

[13] ——, "Optimization based rate control for multirate multicast sessions," in *Proc. IEEE INFOCOM*, Apr. 2001, pp. 123–132.

[14] F. P. Kelly, "Fairness and stability of end-to-end congestion control," *Eur. J. Control*, vol. 9, pp. 159–176, 2003.

[15] 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, no. 3, pp. 237–252, Mar. 1998.

[16] S. Kunniyur and R. Srikant, "End-to-end congestion control: Utility functions, random losses, and ECN marks," *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 689–702, Oct. 2003.

[17] S. H. Low and P. Varaiya, "Dynamic behavior of a class of adaptive routing protocols (IGRP)," in *Proc. IEEE INFOCOM*, Mar. 1993, pp. 610–616.

[18] S. H. Low, "A duality model of TCP and queue management algorithms," *IEEE/ACM Trans. Networking*, vol. 11, no. 4, pp. 525–536, Aug. 2003.

[19] S. H. Low and D. E. Lapsley, "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861–874, Dec. 1999.

[20] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Control Syst. Mag.*, vol. 22, no. 1, pp. 28–43, Feb. 2002.

[21] S. H. Low, L. Peterson, and L. Wang. (2002, Mar.) Understanding Vegas: A duality model. *J. ACM*. [Online], vol (2), pp. 207–235

[22] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.

[23] L. Massoulie and J. Roberts, "Bandwidth sharing: Objectives and algorithms," *IEEE/ACM Trans. Networking*, vol. 10, no. 3, pp. 320–328, Jun. 2002.

[24] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556–567, Oct. 2000.

[25] J. Moy. (1998, Apr.) OSPF Version 2. IETF RFC 2328. [Online]. Available: http://www.faqs.org/rfcs/rfc2328.html

[26] S. Plotkin, D. Shmoys, and E. Tardos, "Fast approximation algorithms for fractional packing and covering problems," *Math. Oper. Res.*, pp. 257–301, 1995.

[27] R. Srikant, *The Mathematics of Internet Congestion Control*. Cambridge, MA: Birkhauser, 2004.

[28] J. Wang, L. Li, S. H. Low, and J. C. Doyle, "Can TCP and shortest-path routing maximize utility," in *Proc. IEEE INFOCOM*, Apr. 2003, pp. 2049–2056.

[29] B. M. Waxman, "Routing of multipoint connections," *IEEE J. Select. Areas Commun.*, vol. 6, no. 9, pp. 1617–1622, Sep. 1988.

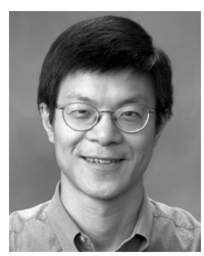
Jiantao Wang (S'05) received the B.S. and M.S. degrees in automatic control from Tsinghua University, Beijing, China, in 1997 and 1999, respectively. He has been a graduate student in Control and Dynamical Systems at the California Institute of Technology, Pasadena, since 1999.

His research interests include network congestion control, active queue management, and routing.

Lun Li (S'05) received the B.S. degrees in optics and automatic control from Tsinghua University, Beijing, China, in 1999 and the M.S. degree in mechanical engineering from the University of California, Berkeley, in 2001. She has been pursuing the Ph.D. degree in electrical engineering at the California Institute of Technology, Pasadena, since 2001.

Her research interests include network congestion control, routing, and Internet topology.

Ms. Li has received the Best Student Paper Award from ACM Sigcomm 2004.



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, Berkeley, both in electrical engineering.

He was with AT&T Bell Laboratories, Murray Hill, NJ, from 1992 to 1996 and with the University of Melbourne, Victoria, Australia, from 1996 to 2000. He is currently an Associate Professor with the California Institute of Technology, Pasadena, and a Senior Fellow of the University of Melbourne. His research interests are in the control and optimization of networks and protocols.

Dr. Low was a co-recipient of the IEEE Bennett Prize Paper Award in 1997 and the 1996 R&D 100 Award. He is on the editorial board of the IEEE/ACM TRANSACTIONS ON NETWORKING, *ACM Computing Surveys*, *NOW* Foundations and Trends in Networking, and the *Computer Networks Journal*, and is a Senior Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.

John C. Doyle received the B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1977, and the Ph.D. degree in mathematics from the University of California, Berkeley, in 1984.

He has served as a consultant to Honeywell Technology Center since 1976, and is the John G. Braun Professor of Control and Dynamical Systems, Electrical Engineering, and Bioengineering at the California Institute of Technology, Pasadena.

Dr. Doyle received the IEEE Centennial Outstanding Young Engineer Award as the top young researcher (under 40) from the IEEE Control Systems Society as part of the IEEE Centennial celebration in 1984. He has received the 1976 IEEE Hickernell Award, the 1983 American Automatic Control Council (AACC) Eckman Award, and the 1984 Bernard Friedman Award. He also received the IEEE Control Systems Award for his fundamental contribution to the analysis and control of uncertain systems in 2004. He has been both an ONR and NSF Presidential Young investigator. His outstanding paper awards include the AACC Hugo Schuck Award for the 1995 American Control Conference, IEEE Transactions George S. Axelby Outstanding Paper Awards (twice), the 1991 IEEE Baker Prize, and the 2004 ACM SIGCOMM best student paper.