

# Which Protocol? Mutual Interaction of Heterogeneous Congestion Controllers

Vinod Ramaswamy, Diganto Choudhury and Srinivas Shakkottai *Member, IEEE*

**Abstract**—A large number of congestion control protocols have been proposed in the last few years, with all having the same purpose—to divide available bandwidth resources among different flows in a fair manner. Each protocol operates on the paradigm of some conception of link price (such as packet losses or packet delays) that determines source transmission rates. Recent work on network utility maximization has brought forth the idea that the fundamental price or Lagrange multiplier for a link is proportional to the queue length at that link, and that different congestion metrics (such as delays or drops) are essentially ways of interpreting such a Lagrange multiplier. We thus ask the following question: Suppose that each flow has a number of congestion control protocols to choose from, which one (or combination) should it choose? We introduce a framework wherein each flow has a utility that depends on throughput, and also has a disutility that is some function of the queue lengths encountered along the route taken. Flows must choose a combination of protocols that would maximize their payoffs. We study both the socially optimal, as well as the selfish cases to determine the loss of system-wide value incurred through selfish decision making, so characterizing the “price of heterogeneity”. We also propose tolling schemes that incentivize flows to choose one of several different virtual networks catering to particular needs, and show that the total system value is greater, hence making a case for the adoption of such virtual networks.

**Index Terms**—Congestion controllers, non-cooperative games, TCP/IP, PMP.

## I. INTRODUCTION

RECENT years have seen the design of a large number of congestion control protocols for use on the Internet. Their designs all revolve around the idea that link congestion is indicated by some notion of “price”, which the source can respond to. Different congestion price metrics include packet loss, packet marks, packet delays or some combination thereof. However, the relative value of one protocol versus another is not well understood. For example, it might be conjectured that a delay sensitive application would consider using a protocol that has a delay-based congestion metric, and a throughput

Manuscript received April 20, 2012; revised December 17, 2012; accepted February 27, 2013; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor C. Dovrolis. This work was supported in part by the NSF under Grants CNS-0904520, CNS-0963818, and CNS-1149458; the Google Faculty Research Awards program under Grants HDTRA1-13-1-0030 and FA9550-13-1-0008; and the Google Faculty Research Awards program.

V. Ramaswamy and S. Shakkottai are with Dept. of ECE, Texas A&M University, College Station, Texas, USA (e-mail: vinod83@tamu.edu, sshakkot@tamu.edu).

D. Choudhury was with Dept. of ECE, Texas A&M University, College Station, Texas, USA. He is now with Maxim Integrated, Dallas, Texas, USA (e-mail: dchoudhury@tamu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>

Digital Object Identifier 10.1109/TNET.2013.2262773

maximizing application might favor a loss-based metric. How should applications choose the protocol to use?

An analytical framework for network resource allocation was developed in seminal work by Kelly *et al.* [1]. If the flow  $i$  has a rate  $x_i \geq 0$  and the utility associated with such a flow is represented by a concave, increasing function  $U_i(x_i)$ , the objective is

$$\max \sum_{i \in \mathcal{N}} U_i(x_i) \quad (1)$$

$$\text{s.t. } y_l \leq c_l, \forall l \in \mathcal{L} \quad (2)$$

where  $\mathcal{N}$  is the set of sources,  $\mathcal{L}$  the set of links,  $c_l$  the capacity of link  $l \in \mathcal{L}$ . Also let  $R$  be the routing matrix with  $R_{li} = 1$  if the route associated with source  $i$  uses link  $l$ . The load on link  $l$  is  $y_l = \sum_{r \in \mathcal{N}} R_{lr} x_r$ . The problem can be solved using ideas based on *Primal-Dual* system dynamics [1]–[5] to yield a set of controllers. At the source we have

$$\text{Source: } \dot{x}_i(t) = \kappa_i \left( U_i'(x_i(t)) - \sum_{l:l \in \mathcal{L}} R_{li} p_l(t) \right)_{x_i}^+ \quad (3)$$

where  $\kappa_i > 0$ , and the notation  $(\phi)_{\xi}^+$  is used to denote the function

$$(\phi)_{\xi}^+ = \begin{cases} \phi & \xi > 0 \\ \max\{\phi, 0\} & \xi = 0. \end{cases} \quad (4)$$

(4) ensures that  $x$  is non-negative. The controller in (3) has an attractive interpretation that the source rate of flow  $i$  responds to feedback in the form of link prices  $p_l(t)$ , with the end-to-end price being calculated as the sum of prices on all links that the flow traverses—something that is common to all congestion control protocols. Source rate is always non-negative, which is enforced by the definition of the function in (4). The price  $p_l(t)$  at link  $l$  is calculated using

$$\text{Link: } \dot{p}_l(t) = \rho(p_l(t)) \left( \sum_{j \in \mathcal{N}} R_{lj} x_j(t) - c_l \right)_{p_l(t)}^+ \quad (5)$$

(5) ensures that the price is non-negative. Each link has a buffer in which packets are queued. If the total load at a link  $l$  given by  $\sum_{j \in \mathcal{N}} R_{lj} x_j(t)$  is greater than the capacity  $c_l$ , the queue length increases, while if it is less than  $c_l$ , the queue length decreases as seen in (5). The queue length is always non-negative, as enforced by the definition in (4). The gain parameter  $\rho(p_l)$  is any positive function. Thus, the *link-price*  $p_l(t)$  can be identified with the *queue length* at link  $l$ . It has been shown [1]–[5] that the above control scheme converges to the optimal solution to the problem in (1).

While this framework indicates that the fundamental price of a link is proportional to queue length, congestion control protocols use several different congestion metrics. For example, TCP Reno [6] uses packet drops (or marks) as its price metric, while TCP Vegas uses end-to-end delay [2]. Other protocols include Scalable TCP [7] (that uses loss-feedback, and allows scaling of rate increases/decreases based on network characteristics), FAST-TCP [8] (that uses delay-feedback, and is meant for high bandwidth environments), and TCP-Illinois [9] (that uses loss and delay signals to attain high throughput). However, drops, marks, and delays are all functions of the queue length. Thus, *a key difference between protocols is their way of interpreting queue length information.*

A fall out of different price-interpretations is that when flows choose distinct congestion control protocols, they do not obtain the same throughput on shared links. For example, studies such as [10]–[13] study inter-protocol as well as intra-protocol fairness, while [14] considers a game of choosing between protocols, assuming that a certain throughput would be guaranteed per combination.

Throughput alone does not fully capture the performance of an application, since it might also be impacted by queueing effects such as delay and packet loss. We consider applications that might have different sensitivities to queueing. Indeed, a large fraction of Internet traffic consists of file transfers (less delay sensitive) and buffered video streams (more delay sensitive) from data centers or content distribution networks. We model these flows as having (possibly different) utilities for throughput, and disutilities for the queueing encountered on their respective paths.

We anticipate for a future Internet architecture where multiple congestion controlling schemes are available to cater the needs of different service classes and the flows are allowed choose the ones according to their service preferences. Hence, we assume that flows play “fair” in that they choose to follow the constraints imposed by employing *some* form of congestion control. Thus, the flows choose from a set of “reasonable” congestion control mechanisms, for example variants of TCP, so as to maximize their payoff that is utility minus disutility.

Our objective is similar to the proposal in [15], where a system design for virtual links tailored for flows that are rate sensitive (R) and delay sensitive (D) is presented. The idea is that an R-flow would pick the virtual link where it is guaranteed higher rate, whereas a D-flow would pick one where it is guaranteed a lower delay. However, unlike that work, we have two basic differences. First, we explicitly model utility (for throughput) and disutility (for queueing) for all kinds of flows, rather than assume that D-type flows would be willing to live with smaller rate. This enables us to explore the space of multiple classes of service with tolling, since it gives an objective measure on the choice made by the flow. Second, we allow a choice between TCP flavors (i.e., interpretation of queue length by congestion controllers) according to the application in question. However, in [15] the only way to reduce delay is to have short buffers for the D service class, which might also result in more losses.

Our finding is that if the number of flows in the system is large, the optimal strategy of a flow is to choose a price

interpretation from among the space of available ones that is most similar to its disutility function. Using this finding, we can characterize the total system value to all flows, and we show that the ratio of this value to the optimum value can be arbitrarily small. Finally, we consider the situation in which we create multiple virtual networks with tolling, with each flow having a choice between networks and between protocols. We show that we can fix the tolls such that the overall system value can be increased significantly, in spite of the toll. We next present our model and summarize our main results.

## II. MODEL AND MAIN RESULTS

We consider a system in which each flow  $i \in \mathcal{N}$  has a so-called  $\alpha$ -fair utility function [16],

$$U_i(x_i) \triangleq w_i x_i^{1-\alpha_i} / (1 - \alpha_i), \quad (6)$$

with  $\alpha_i \geq 1$ , and a disutility that depends on the vector of link prices  $p$  as

$$\tilde{U}_i(x_i, p) \triangleq \sum_{l \in \mathcal{L}} R_{li}(p_l/\tau_i)^\beta x_i, \quad (7)$$

where  $\beta > 1$  is a constant. The overall payoff is the difference of the two, given by

$$F_i(x_i, p) \triangleq U_i(x_i) - \tilde{U}_i(p). \quad (8)$$

The  $\alpha$ -fair utility function was proposed by Mo et. al [16] as a method of capturing a large class of fairness measures based on the value of  $\alpha$  used. For instance, they showed that  $\alpha \rightarrow 1$  results in *proportional* fairness, while  $\alpha \rightarrow \infty$  results in *max-min* fairness. The form of the disutility function is such that based on  $\beta$ , the disutility can be (almost) linear in queue length (which in turn is proportional to delay, weighted by the parameter  $T$ ), to gradually increasing convexity as  $\beta$  rises, to a sharp cutoff for large  $\beta$ . The threshold parameter  $\tau_i$  in (7) models the flow’s sensitivity to queue length, with a small value of  $\tau_i$  indicating high sensitivity (e.g., delay sensitive applications need short queue lengths) and a large value indicating low sensitivity (e.g., loss sensitive applications are affected only by buffer overflow).

We define a set of protocols  $\mathcal{T}$ , with cardinality  $T = |\mathcal{T}|$ . Each protocol  $z \in \mathcal{T}$  is associated with a price-interpretation function  $m^z(p_l) \triangleq (p_l/T_z)^\beta$ . Note that these price-interpretation functions take the same form as disutilities, and model the way in which a particular protocol  $z \in \mathcal{T}$  interprets link prices<sup>1</sup>. Again, a loss-based protocol would have a high value of  $T_z$ , while a delay-based protocol would have a low value. This corresponds to the fact that in a protocol that is modulated by buffer over flows such as TCP Reno, the queue length has no impact until a maximum threshold (buffer size) is reached, after which the price is very high ( $T_z = \text{buffer size}$  here). Similarly, TCP Vegas (approximately) decides on whether the achieved throughput is too high or too low as compared to a threshold, which in turn can be related to a threshold on the per-packet delay seen by the flow ( $T_z$  is less than the buffer size here). Now, while a flow  $i$  cannot change

<sup>1</sup>We will refer to “price-interpretation functions” and “protocols” interchangeably.

its disutility function parameterized by  $\tau_i$  it can choose to use a combination of protocols as it finds appropriate. A particular flow  $i$ 's choice could take the form

$$q_i(p) \triangleq \sum_{z \in \mathcal{T}} \epsilon_i^z \sum_{l=1}^L R_{li} m^z(p_l) \quad (9)$$

where  $\sum_{z \in \mathcal{T}} \epsilon_i^z = 1$ , and  $\epsilon_i^z \geq 0$ . The convex combination models the idea that a flow sometimes measures price in one way (e.g., delay-based) and sometimes in another way (e.g., loss-based).  $\epsilon_i^z$  can be thought of as the probability with which flow  $i$  uses protocol  $z$ . For example, this situation might correspond to a flow using delay and loss measurements simultaneously, and responding to congestion signals (loss or delay) probabilistically. We refer to the choice  $[\epsilon_i^1, \epsilon_i^2, \dots, \epsilon_i^T]$ , made by flow  $i$  as  $\epsilon_i \in \mathbf{E}_i \triangleq \{\epsilon_i : \sum_{z \in \mathcal{T}} \epsilon_i^z = 1, \epsilon_i^z \geq 0\}$ . Further, we denote aggregate choices of all flows by  $\epsilon \in \mathcal{E} \triangleq \prod_{i \in \mathcal{N}} \mathbf{E}_i$ , and will refer to  $\epsilon \in \mathcal{E}$  as a *protocol-profile*.

We first show in Section III that for a given protocol-profile, the bandwidth allocations (and hence the payoffs) are unique. Further, a primal-dual type control will converge to this unique bandwidth allocation. The result is essentially a consistency check that allows us to analytically determine the payoffs as a function of the protocol-profile chosen.

We show in Section IV that all bandwidth allocations that are attainable by a protocol-profile over  $T$  protocols with  $m^1(p) \geq m^2(p) \geq \dots \geq m^T(p)$  are attainable by a protocol-profile over just the two protocols  $m^1(p)$  and  $m^T(p)$ . The result has the appealing interpretation that when  $m^z(p) = (p/T_z)^\beta$ , it is sufficient to only consider the “strictest” interpretation (smallest  $T_z$ , which can be thought of as delay-based feedback) and the most “lenient” (largest  $T_z$ , associated with loss-based feedback). We next show that with two protocols with  $T_s < T_l$ , the bandwidth allocation received by a flow  $i$  is decreasing in the weight it places on the strict protocol. Although the proof is involved, the result is intuitive since a strict protocol would always interpret  $p$  as a larger congestion than the lenient protocol. However, since payoffs are the sum of utility and disutility, it does not follow that all flows would choose the protocol with the higher threshold.

We show in Sections V and VI that in many cases, the total system value is maximized when all flows choose to use only  $m^1(p) = (p/T_s)^\beta$ . On the one hand if flows have price-insensitive payoffs, the protocol-profile used does not matter as long as all of them use the same profile. On the other hand, if there is a mix of flows, some of which have a large disutility function (price-sensitive) and others which do not (price-insensitive), using the strict price-interpretation  $m^1(p) = (p/T_s)^\beta$ , ensures that the price does not become too large for all flows, which maximizes system value.

In Sections V and VI, we also consider the case flows use selfish optimizations to choose their protocol-profiles and study the Nash equilibrium. If all flows have price-insensitive payoffs, then they all choose the lenient price-interpretation  $m^2(p) = (p/T_l)^\beta$ . This case can be mapped to throughput maximizing flows all choosing TCP Reno. If we have a mix of flow types sharing a link, it turns out that the price-sensitive flows with disutility function parametrized by  $\tau \leq T_s$ , choose

the strict price-interpretation  $m^1(p) = (p/T_s)^\beta$ , regardless of the choice of others. Similarly, the price-sensitive flows with disutility threshold  $\tau \geq T_l$ , choose the lenient price-interpretation  $m^2(p) = (p/T_l)^\beta$ . While the other flows may employ mixed strategies. When the number of flows in the system is large, a flow with disutility threshold  $\tau$  picks a mixed strategy that yields an effective price interpretation  $(p/\tau)^\beta$ . The result is interesting since it suggests that a delay sensitive application cannot do any better in terms of overall payoff even if it chooses a more lenient protocol. We also characterize the ratio of system value in the game versus the social optimum for the single-link case to determine an efficiency ratio, which can be quite high.

Finally, in Section VIII we introduce virtual networks, each of which is assigned a certain fraction of the capacity, and chooses a toll. Flows can choose a network and protocols. The idea is similar to Paris Metro Pricing (PMP) [17]–[19], and we show that the system value at Nash equilibrium can be higher overall in spite of tolling. The result suggests that the Internet might benefit by having separate tiers of service for delay-sensitive and loss-sensitive flows.

### III. PROBLEM FORMULATION

We assume that for each link, there exists at least one flow that uses only that link. The assumption implies that all links have a non-zero price. We hypothesize from (3) and (5) that the payoffs should be determined by the protocol-profile  $\epsilon$  as

$$x_i^*(p^*, \epsilon_i) = (U_i')^{-1} \left( \sum_{z=1}^T \epsilon_i^z \sum_{l=1}^L R_{li} m^z(p_l^*) \right), \quad (10)$$

with  $\epsilon_i \in \mathbf{E}_i$  and for all  $l \in \mathcal{L}$ .

$$\sum_{i=1}^N R_{li} x_i^*(p^*, \epsilon_i) = c_l \quad p_l^* > 0, \quad (11)$$

Note that although we have denoted  $x^*$  as depending on both  $\epsilon$  and  $p^*$ , the prices themselves depend on  $\epsilon$  through  $x^*$ , and the solution  $(x^*(\epsilon), p^*(\epsilon))$  (if it exists) is solely a function of  $\epsilon$ . We show that the equilibrium exists, and can be reached using Primal-Dual dynamics. We have the following proposition.

**Proposition 1.** *Given any protocol-profile  $\epsilon$ , Primal-Dual dynamics converge to the unique solution  $(x^*, p^*)$  of the conditions (10) and (11).*

*Proof:* For price-interpretation functions of the form  $(p/T_z)^\beta$ , the source dynamics in (3) can be re-written as

$$\dot{x}_i(t) = \kappa_i \left( U_i'(x_i) - \left( \sum_{z=1}^T \epsilon_i^z \left( \frac{T_1}{T_z} \right)^\beta \right) \sum_{l=1}^L R_{li} m^1(p_l) \right)_{x_i}^+$$

where  $m^1(p_l) = (p_l/T_1)^\beta$ . Let  $U_i(x_i) = \frac{1}{\zeta_i} U_i(x_i)$  where  $\zeta_i = \sum_{z=1}^T \epsilon_i^z (T_1/T_z)^\beta$ , and let  $\kappa_i = \zeta_i$ . Then the above equation can be modified as

$$\dot{x}_i(t) = \zeta_i \left( U_i'(x_i(t)) - \sum_{l=1}^L R_{li} m^1(p_l(t)) \right)_{x_i}^+. \quad (12)$$

Now, in (5) choose  $\rho(p_l) = \frac{1}{m^1(p_l)}$ , where  $m^1$  is derivative of  $m^1$ . Then the price-update equation can be re-written as,

$$\dot{m}^1(p_l(t)) = \left( \sum_{i=1}^N R_{li} x_i(t) - c_l \right)_{p_l}^+. \quad (13)$$

Equations (12) and (13) correspond to the primal-dual dynamics of the following convex maximization problem

$$\begin{aligned} \max_{x > 0} \quad & \sum_{i=1}^N \mathcal{U}_i(x_i) \\ \text{subject to} \quad & \sum_{i=1}^N R_{li} x_i \leq c_l, \quad \forall l \in \mathcal{L}. \end{aligned}$$

The above is a convex optimization problem with a unique solution satisfying (10) and (11). Thus, by the usual Lyapunov argument [2]–[5] Primal-Dual dynamics converge to this solution. Note that our choice of price interpretation makes it a special case of the result in Appendix A Case-1 of [11]. ■

We are now in a position to ask questions about what the flows' payoffs would look like at such an equilibrium, and how this would impact the choice of the protocol-profile. Recall that the payoff obtained by a flow when the system state is at  $x^*(\epsilon), p^*(\epsilon)$  is given by

$$F_i(\epsilon) = U_i(x_i^*(\epsilon)) - \tilde{U}_i(p^*(\epsilon)). \quad (14)$$

We define a system-value function  $V$ , which is equal to the sum of payoff functions of all flows in the network,

$$V(\epsilon) = \sum_{i=1}^N F_i(\epsilon). \quad (15)$$

Our first objective is to find an optimal protocol-profile that maximizes the system-value function.

$$\text{Opt:} \quad \max_{\epsilon \in \mathcal{E}} V(\epsilon). \quad (16)$$

Let  $\epsilon_S^*$  be an optimal profile vector for the above problem. Then we refer to  $V_S = V(\epsilon_S^*)$  as the **value of the social optimum**.

An alternative would be for flows to individually maximize their own payoffs. However, such a proceeding might not lead to an optimal system state that maximizes the value function (15). We characterize the equilibrium state of such a selfish behavior by modeling it as a strategic game.

Let  $\mathcal{G} = \langle \mathcal{N}, \mathcal{E}, \mathcal{F} \rangle$  be a strategic game, where  $\mathcal{N}$  is the set of flows (players),  $\mathcal{E}$  is the set of all protocol profiles (action sets) and  $\mathcal{F} = \{F_1, F_2, \dots, F_N\}$ , where  $F_i : \mathcal{E} \rightarrow \mathbb{R}$  is the payoff function of user  $i$  defined in (14). Define  $\epsilon_{-i} = [\epsilon_1, \epsilon_2, \dots, \epsilon_{i-1}, \epsilon_{i+1}, \epsilon_N]$ , *i.e.*, this represents the choices of all flows except  $i$ . Then  $\epsilon = [\epsilon_i, \epsilon_{-i}]$ . For any fixed  $\epsilon_{-i}$ , flow  $i$  maximizes its payoff as shown below.

$$\text{Game:} \quad \max_{\epsilon_i \in \mathbf{E}_i} F_i(\epsilon_i, \epsilon_{-i}) \quad \forall i \in \mathcal{N}. \quad (17)$$

The game is said to be at a Nash equilibrium when flows do not have any incentive to unilaterally deviate from their current state. We define  $\epsilon_G^*$  as a Nash equilibrium of the game  $\mathcal{G}$  if

$$(\epsilon_G^*)_i = \arg \max_{\epsilon_i \in \mathbf{E}_i} F_i(\epsilon_i, (\epsilon_G^*)_{-i}), \quad \forall i \in \mathcal{N}$$

We refer to  $V_G = V(\epsilon_G^*)$  as the **value of the game**. Finally, we define the ‘‘Efficiency Ratio ( $\eta$ )’’ as

$$\eta = \frac{V_G}{V_S}. \quad (18)$$

#### IV. BASIC RESULTS

We first show that a  $T$ -protocol network can be replaced with an equivalent 2-protocol network. Consider a  $T$ -protocol network with price interpretation functions  $[m^1, m^2, \dots, m^T]$ . Let  $\epsilon \in \mathcal{E}_T$  be a profile state in the  $T$ -network. Then the equilibrium rate vector  $x^*(\epsilon)$  and price vector  $p^*(\epsilon)$  satisfy the equilibrium conditions (10) and (11). Now, consider a 2-protocol network with price interpretation functions  $m^1$  and  $m^T$ . Note that  $m^1 \geq m^z \geq m^T, z = 2, \dots, T-1$ . Let  $\mu \in \mathcal{E}_2$  be a profile state in the 2-protocol network.

**Proposition 2.** *For any equilibrium  $(x^*(\epsilon), p^*(\epsilon))$  in a  $T$ -protocol network,  $\exists$  a protocol-profile  $\mu$  s.t.  $(x^*(\epsilon), p^*(\epsilon))$  is also an equilibrium for the 2-protocol network.*

*Proof:* For any given  $\epsilon \in \mathcal{E}_T$ , let  $(x^*(\epsilon), p^*(\epsilon))$  be an equilibrium pair that satisfies the equilibrium conditions (10) and (11), which are reproduced below for clarity.

$$\begin{aligned} x_i^*(\epsilon) &= (U'_i)^{-1} \left( \sum_{z=1}^T \epsilon_i q_i^{z*} \right), \quad \forall i \in \mathcal{N}, \\ R x^*(\epsilon) &= c, \quad p_l^* > 0, \quad \forall l \in \mathcal{L}. \end{aligned}$$

where  $q_i^{z*} = \sum_{l=1}^L R_{li} m^z(p_l^*(\epsilon))$ . The fact that  $m^T \leq m^z \leq m^1$ , implies,  $q_i^{T*} \leq q_i^{z*} \leq q_i^{1*}, \quad \forall i \in \mathcal{N}, z \in \mathcal{T}$ . Since both  $m^1$  and  $m^T$  are strictly increasing functions, there exists a unique  $\mu_i \in [0, 1]$ , such that,

$$\sum_{z=1}^T \epsilon_i q_i^{z*} = \mu_i q_i^{1*} + (1 - \mu_i) q_i^{T*}.$$

Now, we have

$$\begin{aligned} x_i^*(\epsilon) &= (U'_i)^{-1} \left( \sum_{z=1}^T \epsilon_i q_i^{z*} \right) \\ &= (U'_i)^{-1} (\mu_i q_i^{1*} + (1 - \mu_i) q_i^{T*}), \quad \forall i \in \mathcal{N}, \\ R x^*(\epsilon) &= c, \quad p_l^* > 0, \quad \forall l \in \mathcal{L}. \end{aligned}$$

The above equations correspond to the equilibrium conditions of a 2-protocol network with price interpretation functions  $m^1$  and  $m^T$ . Therefore, there exists a protocol-profile  $\mu = [\mu_1, \dots, \mu_N]$  such that  $(x^*(\epsilon), p^*(\epsilon))$  is an equilibrium pair of 2-protocol network. ■

The above proposition shows that any equilibrium state of a  $T$ -protocol network can be obtained with an equivalent 2-protocol network. Therefore we restrict our study to 2-protocol networks with a ‘‘strict’’ price interpretation  $m^s = (\frac{p}{T_s})^\beta$  and a ‘‘lenient’’ price interpretation  $m^l = (\frac{p}{T_l})^\beta$ , *i.e.*,  $T_s < T_l$ . Also, we redefine the protocol profile of flow  $i$ ,  $\epsilon_i$ , as is  $\epsilon_i = \epsilon_i^1$ , where  $\epsilon_i^1$  is the weight applied on the strict price interpretation. Finally, the equilibrium rate of flow  $i$  can be written in terms

of  $m^s$  and  $m^l$  as follows:

$$\begin{aligned} x_i^*(\epsilon) &= (U_i')^{-1} \left( \sum_{l=1}^L R_{li} (\epsilon_i m^s(p_l^*) + (1 - \epsilon_i) m^l(p_l^*)) \right) \\ &= (U_i')^{-1} \left( (\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta) \sum_{l=1}^L R_{li} m^s(p_l^*) \right). \end{aligned} \quad (19)$$

where  $\epsilon = [\epsilon_1, \epsilon_2, \dots, \epsilon_N]$  is the system protocol-profile. The above result follows from (10).

We next show that the bandwidth allocation received by a flow  $i$  is decreasing in the weight it places on the strict protocol  $m^s(p) = (p/T_s)^\beta$ .

**Proposition 3.** *Let  $x_i^*(\epsilon)$  be the equilibrium rate of flow  $i$  for any  $\epsilon \in \mathcal{E}_2$ . Then,*

$$\frac{\partial x_i^*}{\partial \epsilon_i} \leq 0, \forall i \in \mathcal{N},$$

*Proof:* From (19), we have

$$U_i'(x_i^*) = \sum_{l=1}^L R_{li} m^s(p_l^*) \left( \epsilon_i + (1 - \epsilon_i) \left( \frac{T_s}{T_l} \right)^\beta \right).$$

Then, differentiating above equation with respect to  $\epsilon_j$ , we get,

$$\frac{\partial x_i^*}{\partial \epsilon_j} = A_{ij} + \sum_{l=1}^L \frac{\partial p_l^*}{\partial \epsilon_j} B_{il}, \quad (20)$$

where

$$\begin{aligned} A_{ij} &= \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{l=1}^L R_{li} m^s(p_l^*) \right)}{U_i''(x_i^*)} \delta_{ij}, \quad \text{and} \\ B_{il} &= \frac{R_{li} (m^s)'(p_l^*) (\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta)}{U_i''(x_i^*)}. \end{aligned}$$

Also,  $\delta_{ij} = 1$  if  $i = j$ , and zero otherwise. At equilibrium,  $\sum_{i=1}^N R_{li} x_i^*(\epsilon) = c_l, \forall l \in \mathcal{L}$ . Now, differentiating this equation with respect to  $\epsilon_j$ , we get

$$\sum_{i=1}^N R_{li} \frac{\partial x_i^*}{\partial \epsilon_j} = 0, \quad \forall l \in \mathcal{L}. \quad (21)$$

Replacing  $\frac{\partial x_i^*}{\partial \epsilon_j}$  with (20), we obtain

$$\begin{aligned} &\sum_{i=1}^N R_{li} \frac{(\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta)}{U_i''(x_i^*)} \sum_{k=1}^L R_{ki} (m^s)'(p_k^*) \frac{\partial p_k^*}{\partial \epsilon_j} \\ &+ R_{lj} \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{k=1}^L R_{kj} m^s(p_k^*) \right)}{U_j''(x_j^*)} = 0. \end{aligned}$$

Now, rearranging terms in the above expression, we get,

$$\begin{aligned} &\sum_{k=1}^L (m^s)'(p_k^*) \frac{\partial p_k^*}{\partial \epsilon_j} \sum_{i=1}^N R_{li} R_{ki} \frac{(\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta)}{-U_i''(x_i^*)} \\ &= R_{lj} \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{k=1}^L R_{kj} m^s(p_k^*) \right)}{U_j''(x_j^*)}. \end{aligned}$$

We can represent the above in a matrix form as

$$RWR^T \zeta = r,$$

where

$$\begin{aligned} W &= \text{diag} \frac{(\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta)}{-U_i''(x_i^*)} \\ \zeta &= \left[ (m^s)'(p_1^*) \frac{\partial p_1^*}{\partial \epsilon_j} \quad (m^s)'(p_2^*) \frac{\partial p_2^*}{\partial \epsilon_j} \quad \dots \quad (m^s)'(p_L^*) \frac{\partial p_L^*}{\partial \epsilon_j} \right]^T \\ r &= \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{k=1}^L R_{kj} m^s(p_k^*) \right)}{U_j''(x_j^*)} [R_{1j} \quad \dots \quad R_{Lj}]^T. \end{aligned}$$

Note that  $U_i$  is a strictly concave function and hence  $U_i''(x_i^*) < 0$ . Therefore,  $RWR^T$  is a positive definite matrix. Now, we have

$$\zeta = (RWR^T)^{-1} r. \quad (22)$$

Let  $H = (RWR^T)^{-1}$ , where  $H$  is an  $L \times L$  matrix. Let us represent its elements using  $h_{lm}$ . Thus, from (22), we have

$$\frac{\partial p_l^*}{\partial \epsilon_j} = \frac{\sum_{k=1}^L R_{kj} h_{lk} (1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{k=1}^L R_{kj} m^s(p_k^*) \right)}{(m^s)'(p_l^*) U_j''(x_j^*)}. \quad (23)$$

Let  $V = WR^T(RWR^T)^{-1}R$ . Then, from (20) and (23), we get

$$\frac{\partial x_j^*}{\partial \epsilon_j} = \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{l=1}^L R_{kj} m^s(p_k^*) \right)}{U_j''(x_j^*)} (1 - v_{jj}), \quad (24)$$

$$\frac{\partial x_i^*}{\partial \epsilon_j} = - \frac{(1 - (\frac{T_s}{T_l})^\beta) \left( \sum_{k=1}^L R_{kj} m^s(p_k^*) \right)}{U_j''(x_j^*)} v_{ij}, \quad (25)$$

where  $v_{ij}$  represent elements of  $V$ .

Now, we show that  $\frac{\partial x_j^*}{\partial \epsilon_j}$  is negative given the assumption in the lemma. Note that  $V$  is a projection matrix. The diagonal elements of a projection matrix are positive and less than or equal to unity. i.e.  $v_{jj} \leq 1$ . Then, from (24), we conclude that  $\frac{\partial x_j^*}{\partial \epsilon_j} \leq 0$  and hence have proved the proposition.  $\blacksquare$

The above proposition is intuitive in that a strict protocol would force the flow to cut down its rate for the same price as a lenient protocol.

**Corollary 4.** *In the single link case, the link-price  $p^*$  and the rate vector  $x^*$  satisfies,  $\frac{\partial p^*}{\partial \epsilon_j} < 0$  and  $\frac{\partial x_i^*}{\partial \epsilon_j} > 0$  if  $i \neq j, \forall i, j \in \mathcal{N}$ .*

*Proof:* From (23), (24) and (25), we have

$$\frac{\partial p^*}{\partial \epsilon_j} = \frac{(1 - (\frac{T_s}{T_l})^\beta) m^s(p^*)}{(m^s)'(p^*) U_j''(x_j^*)} \frac{1}{\sum_{r=1}^N \nu_r}, \quad (26)$$

$$\frac{\partial x_i^*}{\partial \epsilon_j} = \frac{(1 - (\frac{T_s}{T_l})^\beta) m^s(p^*)}{U_j''(x_j^*)} \left( \delta_{ij} - \frac{\nu_j}{\sum_{r=1}^N \nu_r} \right), \quad (27)$$

where

$$\nu_i = - \frac{\epsilon_i + (1 - \epsilon_i) (\frac{T_s}{T_l})^\beta}{U_i''(x_i^*)} = \frac{x_i^*}{\alpha_i m^s(p^*)}.$$

The above result follows from (19) and the fact that  $U_i''(x_i^*) = -\frac{\alpha_i}{x_i^*} U_i'(x_i^*)$ . Note that  $U_i''(x) < 0$  since  $U_i$  is strictly concave. Now, the corollary is straightforward from the above results.  $\blacksquare$

Now, we now study different mixes of flow types in order to understand the system value in each case.

## V. FLOWS WITH PRICE-INSENSITIVE PAYOFF

We associate each flow  $i \in \mathcal{N}$  to a class, based on its disutility function of the form  $\sum_{l \in \mathcal{L}} R_{li} (p_l / \tau_i)^\beta x_i$ . We begin by considering a system of flows that have a price-insensitive payoff, *i.e.*,  $\tau_i = \infty \forall i \in \mathcal{N}$ . This means that payoff is solely a function of bandwidth, and we have  $F_i(\epsilon) = U_i(x^*(\epsilon))$ . However, even in this situation, flows must employ congestion control, *i.e.*, they must choose a protocol-profile. From Section (IV), recall that since we only have two protocols, the flow  $i$ 's choice of protocol profile is defined by a scalar value  $\epsilon_i$ . Also note that  $T_z \neq \infty$  for each protocol  $z = 1, 2$ . The system-value is equal to the sum of user payoffs,  $V(\epsilon) = \sum_{i=1}^N U_i(x^*(\epsilon))$ . We then have the following result.

**Proposition 5.** *The system-value is maximized when the protocol choices made by all users are the same. Thus, if  $\epsilon_S^* = \arg \max_{\epsilon \in \mathcal{E}} V(\epsilon)$ , and  $(\epsilon_S^*)_i$  is used to denote the protocol choice made by-profile of user  $i$ , then  $(\epsilon_S^*)_i = (\epsilon_S^*)_j, \forall i, j \in \mathcal{N}$ .*

*Proof:* We first derive an upper bound for system-value  $V(\epsilon)$  and then show that the upper bound is achieved when all sources choose the same protocol. Suppose that  $\mathcal{X} = \{x | Rx = c\}$ . Let  $\hat{x} = \arg \max_{Rx=c} \sum_{i=1}^N U_i(x_i)$ . Note that equilibrium rate  $x^*(\epsilon) \in \mathcal{X}$ , since  $Rx^* = c$ . Then the value of  $\sum_{i=1}^N U_i(x)$  evaluated at  $x^*(\epsilon)$  satisfies

$$V(\epsilon) = \sum_{i=1}^N U_i(x_i^*(\epsilon)) \leq \sum_{i=1}^N U_i(\hat{x}_i).$$

We showed in Proposition 2 that the equilibrium rate  $x^*(\epsilon)$ , is the unique maximizer of the convex problem  $\max_{x>0, Rx=c} \sum_{i=1}^N \frac{1}{\zeta_i} U_i(x_i)$ , where  $\zeta_i = \epsilon_i + (1 - \epsilon_i)(\frac{T_s}{T_l})^\beta$ . Then,  $x^*(\epsilon)$  can be made equal to  $\hat{x}$ , the optimal point in set  $\mathcal{X}$ , by choosing  $\zeta_i = \zeta_j \forall i, j \in \mathcal{N}$ . Such a choice means that

$$\begin{aligned} \zeta_i = \zeta_j &\Rightarrow \epsilon_i + (1 - \epsilon_i)\left(\frac{T_s}{T_l}\right)^\beta = \epsilon_j + (1 - \epsilon_j)\left(\frac{T_s}{T_l}\right)^\beta, \\ &\Rightarrow \epsilon_i = \epsilon_j. \end{aligned}$$

Thus, if  $\epsilon_S^* = \arg \max_{\epsilon \in \mathcal{E}} V(\epsilon) \Rightarrow (\epsilon_S^*)_i = (\epsilon_S^*)_j, \forall i, j \in \mathcal{N}$ . Therefore, the system value is maximized when the protocol choices made by all the users are identical. Also, the maximum value does not depend on the parameters of the selected protocol. ■

We next consider the game in which flows are allowed to choose their protocols selfishly.

**Proposition 6.** *Let  $G = \langle \mathcal{N}, \mathcal{E}, \mathcal{F} \rangle$  be a strategic game with payoff function of user  $i$  is given as  $F_i(\epsilon) = U_i(x_i^*(\epsilon))$ . Then there exists a Nash equilibrium for game  $G$ , and the equilibrium profile for any user  $i \in \mathcal{N}$  is  $(\epsilon_G^*)_i = 0$ .*

*Proof:* Differentiating  $F_i$  w.r.t  $\epsilon_i$ , and using Proposition 3

$$\frac{\partial F_i}{\partial \epsilon_i} = U'(x_i^*(\epsilon)) \frac{\partial x_i^*(\epsilon)}{\partial \epsilon_i} \leq 0$$

Hence,  $F_i(\epsilon)$  is maximized when  $\epsilon_i = 0$ . Therefore,  $(\epsilon_G^*)_i = 0, \forall i \in \mathcal{N}$ . ■

**Efficiency Ratio:** We showed in Proposition 5 that the value function is maximized when all flows pick the same protocol-profile. In Proposition 6 we saw that when each flow selfishly

maximizes its own payoff, there exists a Nash equilibrium under which every source chooses the lowest priced protocol, *i.e.*, the protocol with the higher value of  $T$ . Such a profile is a special case of all flows choosing the same protocol-profile. Thus, value of the social optimum and the value of the game are identical and Efficiency Ratio ( $\eta$ ) is unity.

**Example-1:** Consider the case in which a single link with capacity  $c = 10$  is shared by 2 price-insensitive flows. Users have  $\alpha$ -fair utility functions with  $\alpha = 2$ ,  $w_1 = 100$  and  $w_2 = 100$ . We use price-interpretation functions  $(\frac{p}{2})^2$  and  $(\frac{p}{5})^2$ . Note that the simulation parameters  $\alpha, \beta$  and threshold values are chosen arbitrarily. These parameters may not correspond to any particular protocol used in practice. Nevertheless, the observations made here hold true for any values of  $\alpha \geq 1, \beta > 1$  and  $T_s, T_l, \tau_i > 0$ .

In Figure (1) we show the system value for different choices of protocol profiles. The plot illustrates that system value is maximized when both flows choose the same profile. Figure (2) shows how the payoff function of a flow varies with its protocol profile. We find that regardless of the value of the protocol profile chosen by the other flow, the payoff function is maximized when it picks the lower price protocol.

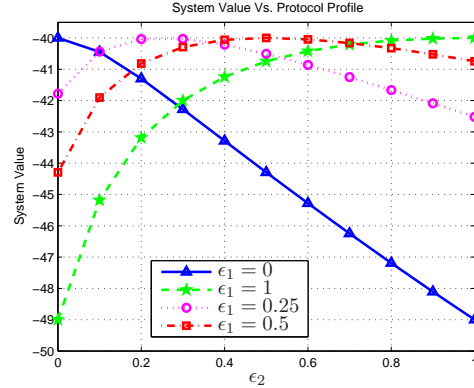


Fig. 1. System Value with price-insensitive flows as a function of the protocol-profile. We observe that the system value is maximized when both flows choose the same protocol-profile.

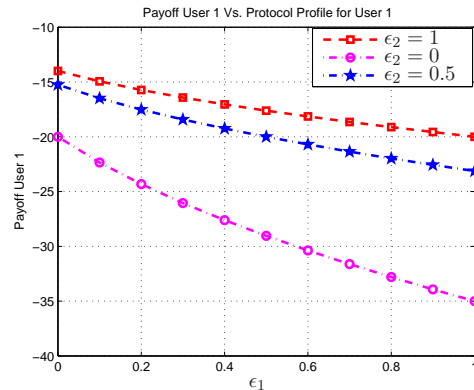


Fig. 2. Payoff of a price-insensitive flow as a function of its protocol-profile. We observe that payoff is maximized when the flow chooses the more lenient price interpretation, regardless of the other flow.

## VI. MIXED ENVIRONMENT

We now consider the case where a network is shared by flows with different disutilities. We identify the optimal

protocol profile that maximizes the system value, and compare it with and the Nash equilibrium. We first study the case of a network consisting of a single link.

### A. Single Link Case

Consider a single link system with capacity  $c$  shared by  $N$  flows. The payoff of user  $i \in \mathcal{N}$  is  $F_i(\epsilon) = U_i(x_i^*(\epsilon)) - \left(\frac{p^*(\epsilon)}{\tau_i}\right)^\beta x_i^*(\epsilon)$ . Then, the system value is  $V(\epsilon) = \sum_{i=1}^N F_i(\epsilon)$ .

**Proposition 7.** *The system-value is maximized when all users pick the protocol with lowest threshold, i.e., if  $\epsilon_S^* = \arg \max_{\epsilon \in \mathcal{E}} V(\epsilon)$ , then  $(\epsilon_S^*)_i = 1, \forall i \in \mathcal{N}$ .*

*Proof:* (Sketch) Recall that  $\alpha_i \geq 1$  by our assumption. Given this assumption, it can be shown through straightforward differentiation that  $\tilde{U}_i(\epsilon_i)$  is a monotonically decreasing function of  $\epsilon_i$ . Now, the value function  $V$  is maximum when  $U(\epsilon)$  is maximized and  $\tilde{U}(\epsilon)$  is minimized. We already know from Proposition 5 that  $U(\epsilon)$  is maximized when all flows choose the same protocol-profile. Coupling this result with the fact that  $\tilde{U}_i(\epsilon_i)$  is decreasing in  $\epsilon_i$ , we see that system value is maximized when  $\epsilon_i = 1, \forall i \in \mathcal{N}$ . ■

We now study the strategic game in which users individually maximize their payoff as in (17). We show that there exists a Nash equilibrium and characterize the protocol-profile.

**Proposition 8.** *Let  $G = \langle \mathcal{N}, \mathcal{E}, \mathcal{F} \rangle$  be a strategic game with payoff of user  $i$  is  $F_i(\epsilon) = U_i(x_i^*(\epsilon)) - \left(\frac{p^*(\epsilon)}{\tau_i}\right)^\beta x_i^*(\epsilon)$ . Then there exists a Nash equilibrium (NE) for Game  $G$ . At NE, flows with greatest sensitivity to price choose the strict protocol, i.e., if  $\tau_i = T_s$ , then  $\epsilon_i = 1$ .*

*Proof:* We will show that  $F_i(\epsilon)$  is quasi-concave, and use the Theorem of Nash to show existence of a NE. Differentiating  $F_i$  w.r.t  $\epsilon_i$ ,

$$\frac{\partial F_i}{\partial \epsilon_i} = (U_i'(x_i^*) - d_i(p^*)) \frac{\partial x_i^*}{\partial \epsilon_i} - d_i'(p^*) x_i^* \frac{\partial p^*}{\partial \epsilon_i}, \quad (28)$$

where  $d_i(p^*) = \left(\frac{p^*}{\tau_i}\right)^\beta$  and  $d_i'(p^*)$  is its derivative. Now, substituting the results from (26) and (27), in the above equation, we get

$$\frac{\partial F_i}{\partial \epsilon_i} = B(U_i'(x_i^*) - d_i(p^*)) \left(1 - \frac{\nu_j}{\sum_{r=1}^N \nu_r}\right) \quad (29)$$

$$- B \frac{d_i'(p^*) x_i^*}{(m^s)'(p^*) \sum_{r=1}^N \nu_r}, \quad (30)$$

where  $B = \frac{(1 - (\frac{T_s}{T_l})^\beta) m^s(p^*)}{U_i''(x_i^*)}$  and  $\nu_i = \frac{x_i^*}{\alpha_i m^s(p^*)}$ . Note that  $B < 0$  since  $U_i''$  is a negative function.

From (19) along with the definitions of  $\nu_i$  and  $d_i(p^*)$ , the above expression can be simplified as follows:

$$\begin{aligned} \frac{\partial F_i}{\partial \epsilon_i} = & \frac{B m^s(p^*) \sum_{r=1, r \neq i}^N \frac{x_r^*}{\alpha_r}}{\sum_{r=1}^N \frac{x_r^*}{\alpha_r}} \left( \epsilon_i + (1 - \epsilon_i) \left(\frac{T_s}{T_l}\right)^\beta - \left(\frac{T_s}{\tau_i}\right)^\beta \right) \\ & - \frac{B m^s(p^*)}{\sum_{r=1}^N \frac{x_r^*}{\alpha_r}} \left(\frac{T_s}{\tau_i}\right)^\beta x_i^*. \end{aligned} \quad (31)$$

We show that if the above expression has a root, then it is unique. The roots are characterized by

$$\epsilon_i + (1 - \epsilon_i) \left(\frac{T_s}{T_l}\right)^\beta = \left(\frac{T_s}{\tau_i}\right)^\beta \left(1 + \frac{x_i^*}{\sum_{r=1, r \neq i}^N \frac{x_r^*}{\alpha_r}}\right). \quad (32)$$

First observe that the left side of the above expression is strictly increasing in  $\epsilon_i$  (since  $T_s < T_l$ ). Since  $\frac{\partial x_i^*}{\partial \epsilon_i} < 0$  and  $\frac{\partial x_r^*}{\partial \epsilon_i} > 0$  if  $r \neq i$  (from Proposition 3 and Corollary 4), the right side of the above expression is strictly decreasing. Therefore, the set of roots of the equation,  $\frac{\partial F_i}{\partial \epsilon_i}(x) = 0$  is a singleton or null set. Thus,  $F_i$  is unimodal or monotonic in  $\epsilon_i$  for any fixed  $\epsilon_{-i}$  and hence quasi concave.

Since  $\epsilon_i \in [0, 1]$  is a non-empty compact convex set, by the theorem of Nash, the quasi concavity of  $F_i(\epsilon_i, \epsilon_{-i})$  guarantees that there exists a  $\epsilon_G^*$ , such that for all  $i = 1, \dots, N$ ,

$$(\epsilon_G^*)_i = \arg \max_{\epsilon_i \in [0, 1]} F_i(\epsilon_i, (\epsilon_G^*)_{-i}).$$

Hence, the first part of the proof is complete.

Now, consider a flow with disutility (per unit rate)  $\left(\frac{p}{\tau_i}\right)^\beta$ , where  $\tau = T_s$ . Replacing  $\tau_i$  with  $T_s$  in (31), we observe that  $\frac{\partial F_i}{\partial \epsilon_i} > 0$  (Note that  $B < 0$ ). Therefore, payoff is maximized when  $\epsilon_i = 1$ . ■

In the next section, we study the characteristics of the NE and show that it is unique.

### B. Nash equilibrium characteristics

We have established the existence of NE of the strategic game (17) in the previous section. We conduct further studies on the properties of NE in this section. First, we derive conditions for the NE system protocol profile. Then, in Proposition 9, we show that the game has a unique NE. Finally, in Proposition 10, we derive the NE strategies of flows when there are large number of flows in the system.

Let  $\hat{\epsilon}$  be a Nash equilibrium system protocol profile (action profile). Then, by definition, it must satisfy the condition that

$$\hat{\epsilon}_i = \arg \max_{\epsilon_i \in [0, 1]} F_i(\epsilon_i, (\hat{\epsilon})_{-i}), \forall i \in \mathcal{N}.$$

Then, from the first order optimality condition, we have

$$\frac{\partial F_i(\hat{\epsilon})}{\partial \epsilon_i} (\epsilon_i - \hat{\epsilon}_i) \leq 0.$$

Consequently, from (31), we get that,  $\forall i \in \mathcal{N}$ ,

$$\gamma(\hat{\epsilon}_i) = \left(\frac{1}{T_s^\beta} \wedge \frac{1}{T_l^\beta} \left(1 + \frac{x_i^*(\hat{\epsilon})}{\sum_{r=1, r \neq i}^N \frac{x_r^*(\hat{\epsilon})}{\alpha_r}}\right)\right) \vee \frac{1}{T_l^\beta}. \quad (33)$$

where  $a \wedge b = \min\{a, b\}$ ,  $a \vee b = \max\{a, b\}$  and  $\gamma(\epsilon_i) = \epsilon_i \left(\frac{1}{T_s}\right)^\beta + (1 - \epsilon_i) \left(\frac{1}{T_l}\right)^\beta$ . In addition, the Nash equilibrium profile must also satisfy,

$$x_i^*(\hat{\epsilon}) = \left(\frac{w_i}{\gamma(\hat{\epsilon}_i)(p^*)^\beta}\right)^{\frac{1}{\alpha_i}}, \quad (34)$$

$$\sum_{i=1}^N x_i^*(\hat{\epsilon}) = c. \quad (35)$$

Here, (34) follows from (19) and the definition of  $U_i(x)$ . Also, (35) follows from the assumption that every link has one flow

using that link alone. Now, we show that the set of Nash equilibria, characterized by (33)-(35), is singleton.

**Proposition 9.** *The strategic game,  $G = \langle \mathcal{N}, \mathcal{E}, \mathcal{F} \rangle$ , has a unique Nash equilibrium.*

*Proof:* To prove by contradiction, assume multiple Nash equilibria exist. Let two distinct NE system protocol profiles be  $\hat{\epsilon}^1$  and  $\hat{\epsilon}^2$ . Also, let  $x_i^1 = x_i^*(\hat{\epsilon}^1)$ ,  $x_i^2 = x_i^*(\hat{\epsilon}^2)$ ,  $p^1 = p^*(\hat{\epsilon}^1)$ ,  $p^2 = p^*(\hat{\epsilon}^2)$ ,  $\gamma_i^1 = \gamma(\hat{\epsilon}^1)$  and  $\gamma_i^2 = \gamma(\hat{\epsilon}^2)$ . Then, by reordering the flow indices, we get that, for some  $k \in \{0, 1, \dots, N\}$ ,

$$\gamma_i^1 > \gamma_i^2 \quad \text{for } i = 1, 2, \dots, k, \quad (36)$$

$$\gamma_i^1 \leq \gamma_i^2 \quad \text{for } i = k + 1, \dots, N. \quad (37)$$

Also, if  $k = 0$ , there exist a flow  $i \in \mathcal{N}$  such that  $\gamma_i^1 < \gamma_i^2$ . We show that the above condition are infeasible for all values of  $k$ , under the NE conditions given by (33-34).

Initially, consider the case when  $k = N$ . Then, from (33), for  $i = 1, 2, \dots, N$ , we have

$$\frac{x_i^1}{\sum_{r=1, r \neq i}^N \frac{x_r^1}{\alpha_r}} > \frac{x_i^2}{\sum_{r=1, r \neq i}^N \frac{x_r^2}{\alpha_r}} \Rightarrow \frac{x_i^1}{\sum_{r=1}^N \frac{x_r^1}{\alpha_r}} > \frac{x_i^2}{\sum_{r=1}^N \frac{x_r^2}{\alpha_r}} \quad (38)$$

$$\Rightarrow \frac{\sum_{r=1}^N \frac{x_r^1}{\alpha_r}}{\sum_{r=1}^N \frac{x_r^1}{\alpha_r}} > \frac{\sum_{r=1}^N \frac{x_r^2}{\alpha_r}}{\sum_{r=1}^N \frac{x_r^2}{\alpha_r}} \quad (39)$$

which is a contradiction. Hence, this case is not feasible. Similarly, we can show that the case when  $k = 0$  is also not feasible.

Now, consider the case when  $1 \leq k < N$ . Also, suppose that  $p^1 \geq p^2$ . Then, from (34), we have

$$x_i^1 < x_i^2, \quad \text{for } i = 1, 2, \dots, k.$$

Let

$$i^* = \arg \max_i \frac{x_i^1}{x_i^2}.$$

Note that  $i^* > k$  and hence,  $\gamma_{i^*}^1 \leq \gamma_{i^*}^2$ . Also, from (35), note that  $x_{i^*}^1 > x_{i^*}^2$ .

Observe that,

$$\frac{x_i^1}{x_{i^*}^1} = \frac{x_i^1 x_{i^*}^2}{x_{i^*}^1 x_i^2} \leq \frac{x_i^2}{x_{i^*}^2},$$

and strict inequality holds if  $i \leq k$ . It follows from the above result that,

$$\frac{x_{i^*}^1}{\sum_{r=1}^N \frac{x_r^1}{\alpha_r}} > \frac{x_{i^*}^2}{\sum_{r=1}^N \frac{x_r^2}{\alpha_r}} \Rightarrow \frac{x_{i^*}^1}{\sum_{r=1, r \neq i^*}^N \frac{x_r^1}{\alpha_r}} > \frac{x_{i^*}^2}{\sum_{r=1, r \neq i^*}^N \frac{x_r^2}{\alpha_r}}.$$

Finally, from (33) and the above result, we get  $\gamma_{i^*}^1 \geq \gamma_{i^*}^2$ . But, from the definition of  $i^*$ , we know that  $\gamma_{i^*}^1 \neq \gamma_{i^*}^2$ . In case  $\gamma_{i^*}^1 = \gamma_{i^*}^2$ , then, from (34) and the assumption that  $p^1 \geq p^2$ , we get  $x_{i^*}^1 \leq x_{i^*}^2$ , which also raises a contradiction. Hence, this case is also not feasible. In similar fashion, we can show that the case in which  $p^1 < p^2$  is also not feasible.

Hence, our assumption that multiple NE exist is not true. Therefore, NE is unique. ■

Next, we characterize the NE in the asymptotic regime.

**Proposition 10.** *When the number of flows in the system,  $N$ , is large, the protocol profile of flow  $i$  at NE,  $\hat{\epsilon}_i$ , satisfies*

$$\hat{\epsilon}_i \left( \frac{1}{T_s} \right)^\beta + (1 - \hat{\epsilon}_i) \left( \frac{1}{T_l} \right)^\beta = \left( \left( \frac{1}{T_s} \right)^\beta \wedge \left( \frac{1}{T_l} \right)^\beta \right) \vee \left( \frac{1}{T_l} \right)^\beta.$$

*Proof:* Recall from (33) that, the NE protocol profile of flow  $i$ , satisfies,

$$\gamma(\hat{\epsilon}_i) = \left( \frac{1}{T_s} \wedge \left( \frac{1}{T_l} \right)^\beta \left( 1 + \frac{x_i^*(\hat{\epsilon})}{\sum_{r=1, r \neq i}^N \frac{x_r^*(\hat{\epsilon})}{\alpha_r}} \right) \right) \vee \frac{1}{T_l}.$$

In order to prove the proposition, we claim that,

$$\lim_{N \rightarrow \infty} \frac{x_i^*(\hat{\epsilon}_i)}{\sum_{r=1, r \neq i}^N \frac{x_r^*(\hat{\epsilon}_i)}{\alpha_r}} = 0, \quad (40)$$

holds true. Before proving the above result, we introduce a few notations: Let  $\alpha_{max} = \max_i \alpha_i$ ,  $\alpha_{min} = \min_i \alpha_i$ ,  $w_{max} = \max_i w_i$  and  $w_{min} = \min_i w_i$ .

Now, the proof of the claim (40) is as follows: From (35), we can show that,

$$\frac{x_i^*(\hat{\epsilon})}{\sum_{r=1, r \neq i}^N \frac{x_r^*(\hat{\epsilon})}{\alpha_r}} \leq \frac{\alpha_{max}}{\frac{c}{x_i^*(\hat{\epsilon})} - 1}.$$

Also, from (34), we have,

$$x_i^*(\hat{\epsilon}) = \left( \frac{w_i}{\gamma(\hat{\epsilon}_i)(p^*(\hat{\epsilon}))^\beta} \right)^{\frac{1}{\alpha_i}} \leq \left( \frac{w_{max} T_l^\beta}{(p^*(\hat{\epsilon}))^\beta} \right)^{\frac{1}{\alpha_{min}}}. \quad (41)$$

The above result follows from the fact that  $\gamma(\hat{\epsilon}_i) \geq \left( \frac{1}{T_l} \right)^\beta$ .

From Corollary 4, we observe that the link-price is a decreasing function of protocol profile of each flow and hence, the system protocol profile  $\epsilon$ . Therefore, the link price achieves the lowest value, when every flow adopts the strict protocol. Then, from (34) and (35), it is easy to show that

$$(p^*(\hat{\epsilon}))^\beta \geq w_{min} \left( \frac{N^{\alpha_{min}}}{c^{\alpha_{max}}} \right) T_s^\beta. \quad (42)$$

Finally, from (41) and (42), we have

$$\frac{x_i^*}{\sum_{r=1, r \neq i}^N \frac{x_r^*}{\alpha_r}} \leq \frac{\alpha_{max}}{\frac{c}{x_i^*} - 1} \leq \frac{\alpha_{max}}{NK - 1}$$

where  $K$  is a constant. The upper bound in the above expression goes to zero for large values of  $N$ . Therefore, the claim in (40) holds true and hence, the proof is completed. ■

**Example-2:** We consider a link with capacity  $c = 10$  shared by two flows with disutilities  $\left(\frac{p}{2}\right)^2$  and  $\left(\frac{p}{5}\right)^2$ , respectively, and  $w_1 = w_2 = 1$ . The other parameters are unchanged from Example-1. We show the system value for different choices of protocol-profiles in Figure 3. The value is maximized when both flows choose the strict protocol. Figure (4) shows how the payoff of each flow varies with its choice of protocol profile, given other's is fixed. We find that for the first (sensitive) flow, the payoff function is maximized when it chooses the strict protocol, regardless of the other flow. But the payoff of the second (less-sensitive) flow is maximized for some combination of protocols. The results validate our findings.

**Example-3:** We consider a link with capacity  $c = 1000$ . There are 40 flows sharing the link. The strict and lenient



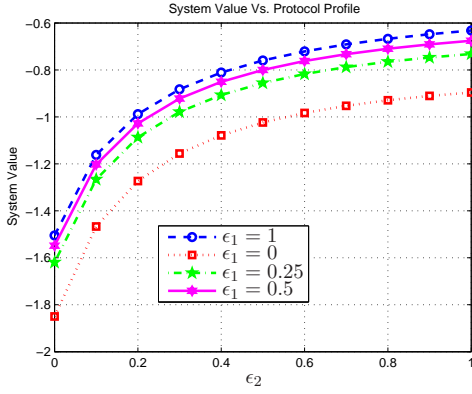


Fig. 3. System value against protocol choices ( $\epsilon_i$ ): Two flows sharing a link.

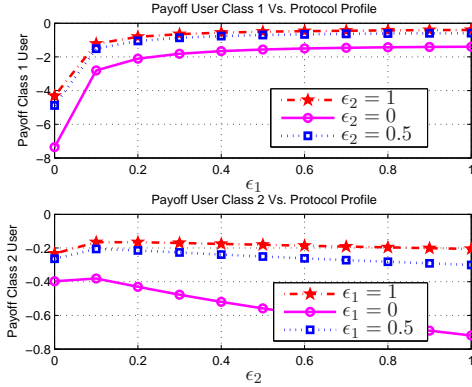


Fig. 4. Payoff against protocol choice ( $\epsilon_i$ ): Two flows sharing a link.

thresholds are  $T_s = 2$  and  $T_l = 7$  respectively. In our simulations, we have set  $\beta = 2$ ,  $\alpha = 2$  for half of users and  $\alpha = 3$  for the other half. There are 10 classes of flows, with each class containing 4 flows. The disutility threshold of a Class  $i$  flow, given by  $\tau_i$ , is chosen according to the following relation:  $(\frac{1}{\tau_i})^\beta = (\frac{1}{T_l})^\beta + ((\frac{1}{T_s})^\beta - (\frac{1}{T_l})^\beta)(i/10)$ .

We choose a candidate flow that belongs to Class 4. We assume that every other flow has chosen their NE protocol profile. That means, the effective price interpretation of a flow belonging to Class  $i$  is  $(\frac{p}{\tau_i})^\beta$ . Figure (5) plots the payoff of the candidate flow as a function of its effective protocol choice  $\gamma(\epsilon_4) = \epsilon_4(1/T_s)^\beta + (1 - \epsilon_4)(1/T_l)^\beta$ , where  $\epsilon_4$  is its protocol profile. As claimed by Proposition 10, the payoff is maximized when  $\gamma(\epsilon_4) = (\frac{1}{\tau_4})^\beta = 0.17$ .

### C. Network Case

We consider a system of flows with log utility functions, which is a special class of an  $\alpha$ -fair utility function with  $\alpha \rightarrow 1$ . The payoff of flow  $i \in \mathcal{N}$  is  $F_i(\epsilon) = w_i \log(x_i^*(\epsilon)) - \sum_{l=1}^L R_{li} (\frac{p_l^*(\epsilon)}{\tau_i})^\beta x_i^*(\epsilon)$ . Then the system-value is  $V(\epsilon) = \sum_{i=1}^N F_i(\epsilon)$ .

**Proposition 11.** *The System-Value function is maximized when all flows pick the higher priced protocol, namely  $m^1 = (\frac{p}{T_s})^\beta$ . Let  $\epsilon_S^* = \arg \max_{\epsilon} V(\epsilon)$ , then  $(\epsilon_S^*)_i = 1, \forall i = 1, \dots, N$ ,*

*Proof:* We can show through straightforward differentiation that the disutility function,  $\tilde{U}_i(\epsilon_i)$ , is a monotonically

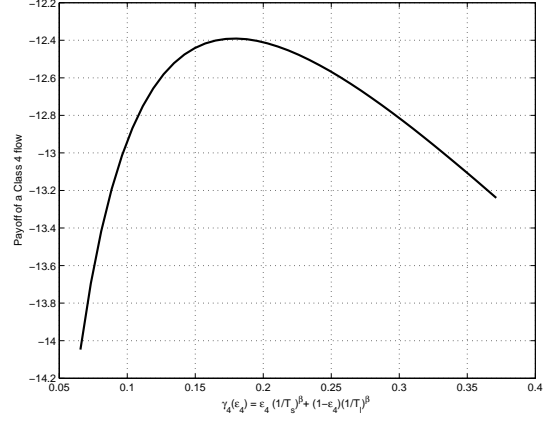


Fig. 5. Payoff of a Class 4 flow is maximized when  $\gamma(\epsilon_4) = (\frac{1}{\tau_4})^\beta = 0.17$ .

decreasing function of  $\epsilon_i$ . The rest of the proof is similar to that of Proposition 7. ■

We now consider a game with two types of flows: price-insensitive flows with zero disutilities, and price-sensitive flows with disutility (per unit rate)  $(\frac{p^*}{T_s})^\beta$ . In this special case, there exists a unique Nash equilibrium. In Proposition 6 we saw that price-insensitive flows pick the lenient protocol at Nash equilibrium irrespective of the choices of the other players. We will now show that price-sensitive flows pick the strict protocol at Nash equilibrium.

**Proposition 12.** *At the unique Nash equilibrium, any flow  $i$  with disutility (per rate)  $(p/T_s)^\beta$  (i.e.  $\tau_i = T_s$ ) picks  $\epsilon_i = 1$ .*

*Proof:* It can be shown through straightforward differentiation that  $\frac{\partial F_i}{\partial \epsilon_i} > 0$  for any flow  $i \in \mathcal{N}$  with disutility (per rate)  $(p/T_s)^\beta$ , which completes the proof. ■

## VII. EFFICIENCY RATIO

We now characterize the loss of system value at Nash equilibrium, as compared to the value of the social optimum. We focus on the case of a single link with capacity  $c$ .

**Proposition 13.** *Assume  $\alpha_i > 1, \forall i \in \mathcal{N}$ . When the number of flows in the system is large,*

$$\eta = \frac{V_G}{V_S} < \hat{\alpha} \left(\frac{T_l}{T_s}\right)^\beta.$$

where  $\hat{\alpha} = \max_i \alpha_i$ .

*Proof:* Let  $\epsilon^* = [\epsilon_1^*, \epsilon_2^*, \dots, \epsilon_N^*]$  be the system protocol profile at social optimum. From Proposition 7, every user chooses the strict protocol at social optimum, i.e.  $\epsilon_i^* = 1, \forall i$ . Hence, from (19), and the definition of  $U_i$ , we have

$$x_i^*(\epsilon^*) = \left( w_i \left( \frac{T_s}{p^*(\epsilon^*)} \right)^\beta \right)^{\frac{1}{\alpha_i}}, \quad \sum_i x_i^*(\epsilon^*) = c. \quad (43)$$

Interpreting  $\left( \frac{p^*(\epsilon^*)}{T_s} \right)^\beta$  as the dual variable, the above equations can be identified as the KKT conditions of the optimization problem given below:

$$\max_x \sum_i \frac{w_i x_i^{1-\alpha_i}}{1-\alpha_i}, \quad \text{subject to} \quad \sum_i x_i = c.$$

And,  $x^*(\epsilon^*)$  is the unique maximizer of the above problem. The payoff of a flow at social optimum, from (8) and the above results, is given by

$$F_i(\epsilon^*) = U_i(x_i^*(\hat{\epsilon}^*)) \left(1 + \mathbf{1}_i(\alpha_i - 1) \left(\frac{T_s}{\tau_i}\right)^\beta\right). \quad (44)$$

where  $\mathbf{1}_i = 1$  if flow  $i$  is a price sensitive flow and zero otherwise. The system value at social optimum is  $V_S = \sum_i F_i(\epsilon^*)$ .

Now, let  $\hat{\epsilon} = [\hat{\epsilon}_1, \hat{\epsilon}_2, \dots, \hat{\epsilon}_N]$  be the system protocol profile at Nash equilibrium. From Proposition 10, equation (19) and the definition of  $U_i$ , we have

$$x_i^*(\hat{\epsilon}) = \left( w_i \left( \frac{T_l \wedge (\tau_i \vee T_s)}{p^*(\hat{\epsilon})} \right)^\beta \right)^{\frac{1}{\alpha_i}}, \quad \sum_i x_i^*(\hat{\epsilon}) = c. \quad (45)$$

Recall that  $a \wedge b = \min\{a, b\}$ ,  $a \vee b = \max\{a, b\}$ . Interpreting  $\left(\frac{p^*(\hat{\epsilon})}{T_s}\right)^\beta$  as the dual variable, the above equations can be identified as the KKT conditions of the optimization problem given below:

$$\max_x \sum_i \frac{w_i \left(\frac{T_l \wedge (\tau_i \vee T_s)}{T_s}\right)^\beta x_i^{1-\alpha_i}}{1 - \alpha_i}, \quad \text{subject to} \quad \sum_i x_i = c.$$

Also,  $x^*(\hat{\epsilon})$  is the unique maximizer of the above problem. Finally, the payoff of a flow is

$$F_i(\hat{\epsilon}) = U_i(x_i^*(\hat{\epsilon})) \left(1 + \mathbf{1}_i(\alpha_i - 1) \left(\frac{T_l \wedge (\tau_i \vee T_s)}{\tau_i}\right)^\beta\right). \quad (46)$$

The system value at NE is  $V_G = \sum_i F_i(\hat{\epsilon})$ .

Now, from the above results and the fact that  $U_i$ 's are negative, since  $\alpha_i > 1$  by the assumption of this proposition, we can show that

$$\begin{aligned} V_G &\geq \hat{\alpha} \sum_i \left(\frac{\tilde{T}}{T_s}\right)^\beta U_i(x_i^*(\hat{\epsilon})) \geq \hat{\alpha} \sum_i \left(\frac{\tilde{T}}{T_s}\right)^\beta U_i(x_i^*(\epsilon_i^*)) \\ &> \hat{\alpha} \left(\frac{T_l}{T_s}\right)^\beta \sum_i U_i(x_i^*(\epsilon_i^*)) \left(1 + \mathbf{1}_i(\alpha_i - 1) \left(\frac{T_s}{T_l}\right)^\beta\right) \\ &= \hat{\alpha} \left(\frac{T_l}{T_s}\right)^\beta V_S, \end{aligned} \quad (47)$$

where  $\hat{\alpha} = \max_i \alpha_i$  and  $\tilde{T} = T_l \wedge (\tau_i \vee T_s)$ . Since  $V_G$  and  $V_S$  are negative, the efficiency ratio  $\eta$ , can be bounded as

$$\eta = \frac{V_G}{V_S} < \hat{\alpha} \left(\frac{T_l}{T_s}\right)^\beta,$$

which completes the proof.  $\blacksquare$

**Example-4:** The exact expression for efficiency ratio is derived for the following special case: We assume that every flow has the same utility function, i.e., in (6),  $w_i = w$  and  $\alpha_i = \alpha, \forall i \in \mathcal{N}$ . We associate the flows, having disutility functions of the form  $\left(\frac{p}{\tau_j}\right)^\beta x$  with Class- $j$ . Assume that there are  $J - 1$  such classes with  $\tau_1 < \tau_2 < \dots < \tau_{J-1}$  and  $\tau_j \in [T_l, T_s], \forall j$ . The flows having zero disutility function is classified as Class  $J$ . For algebraic convenience, we define  $\tau_j = \infty$ . Let  $N_i$  be the number of flows belonging to Class  $i$

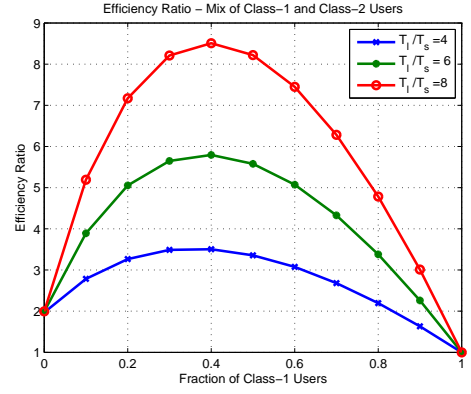


Fig. 6. Efficiency Ratio ( $\eta$ ) in the single link case, plotted against the fraction of Class-1 flows for different ratios of  $T_l/T_s$ . Since  $V_S$  and  $V_G$  were negative in this example, a higher ratio is worse.

and  $n_i = N_i/N$ . Then, the Value of social optimum ( $V_S$ ) and value of game equilibrium ( $V_G$ ) are given by

$$V_S = \frac{N}{1 - \alpha} \left(\frac{c}{N}\right)^{1-\alpha} \sum_{j=1}^J n_j (1 + \mathbf{1}_j(\alpha - 1) \left(\frac{T_s}{\tau_j}\right)^\beta), \quad (48)$$

and

$$V_G = \frac{N \left(\frac{c}{N}\right)^{1-\alpha} S_1}{(1 - \alpha) S_2}, \quad (49)$$

respectively, where

$$S_1 = \left( \alpha \sum_{j=1}^{J-1} n_j \left(\frac{\tau_j}{T_s}\right)^{\left(\frac{\beta}{\alpha}\right)(1-\alpha)} + n_J \left(\frac{T_l}{T_s}\right)^{\left(\frac{\beta}{\alpha}\right)(1-\alpha)} \right)$$

and

$$S_2 = \left( \sum_{j=1}^{J-1} n_j \left(\frac{\tau_j}{T_s}\right)^{\frac{\beta}{\alpha}} + n_J \left(\frac{T_l}{T_s}\right)^{\frac{\beta}{\alpha}} \right)^{1-\alpha}.$$

Also,  $\mathbf{1}_j = 0$  when  $j = J$  and one otherwise. The efficiency ratio,  $\eta$ , is given by

$$\eta = \frac{S_1}{S_2 \sum_{j=1}^J n_j (1 + \mathbf{1}_j(\alpha - 1) \left(\frac{T_s}{\tau_j}\right)^\beta)}. \quad (50)$$

Now, we plot the efficiency ratio for the following case. Let two classes of flows, namely Class 1 and Class 2, are sharing a link. Also, let their disutility thresholds be  $\tau_1 = T_s$  and  $\tau_2 = T_l$  respectively. Letting  $\alpha = 2$  and  $\beta = 3$ , we plot the efficiency ratio ( $\eta$ ), given by (50), in Figure 6. The Figure 6 shows that  $\eta$  increases with  $\left(\frac{T_l}{T_s}\right)$ . Note that a higher ratio is worse. Hence, the performance deteriorates with  $\left(\frac{T_l}{T_s}\right)$ .

## VIII. PARIS METRO PRICING

We have shown in the previous section that when the flows selfishly choose protocols to maximize their own payoff, the system performance at the resulting equilibrium, compared to the socially optimal case, can be much worse. This is due to the fact that, as shown by Proposition 10, the flows with relatively lower disutility functions choose relatively lenient protocols, and hence capture a larger fraction of channel bandwidth leaving not enough for the ones with larger disutility functions

who choose stricter protocols. As a solution to the aforementioned problem, we propose a scheme in which the network is partitioned into virtual subnetworks each having its own queuing buffer, independent price (queue-length) dynamics and fixed entrance toll. A flow is free to choose a protocol along with a subnetwork so as to maximize his own payoff. This scheme is similar to *Paris Metro Pricing* (PMP) [17]. We show that the efficiency of this scheme is superior to the conventional, untolled, single network scheme.

We characterize the performance of the proposed scheme in a single link case. The single link, with capacity  $c$  (bits/sec), is partitioned into  $J$  virtual subnetworks. Let  $S_j$  represent the  $j^{\text{th}}$  sub-network. The bandwidth and toll associated with  $S_j$  are denoted by  $c_j$  and  $\lambda_j$  respectively. Also, let  $\mathbf{c} = [c_1, \dots, c_J]$  and  $\lambda = [\lambda_1, \dots, \lambda_J]$ . We refer to  $\mathbf{c}$  and  $\lambda$  as bandwidth vector and toll vector respectively.

We assume that every flow has the same utility function, i.e, in (6),  $w_i = w$  and  $\alpha_i = \alpha, \forall i \in \mathcal{N}$ . We associate the flows having disutility functions of the form  $(\frac{p}{\tau_j})^\beta x$  to Class- $j$ . We assume that there are  $J-1$  such classes and  $\tau_1 < \tau_2 < \dots < \tau_{J-1}$  with  $\tau_j \in [T_s, T_l]$ . The price insensitive flows are classified as Class- $J$ . For algebraic convenience, we define  $\tau_J = \infty$ . We also assume that there are a large number of flows in each class. Let  $N_j$  represent the number of flows in Class- $j$ .

A flow that seeks to maximize its payoff picks a subnetwork that yields the maximum payoff. Thus, if  $\hat{k}$  is the subnetwork chosen by flow  $i$ ,

$$\hat{k} = \arg \max_{k \in \{1, \dots, J\}} F_{jk} \quad j = 1 \dots, J$$

where  $F_{jk}$  is the payoff of a Class- $j$  flow in  $S_k$ . A Nash equilibrium (NE) here is a state from which none of the flows has an incentive to deviate from its current choice of subnetwork. Note that we already know the flow's choices of protocols in each network so no deviations in protocol are possible. The desired NE is one in which all Class- $j$  flows select  $S_j$ , i.e

$$F_{jj} \geq F_{jk}, \quad \forall j, \forall k. \quad (51)$$

Note that the payoffs received are uniquely determined by the PMP system parameters  $\mathbf{c}$  and  $\lambda$ . Now, we derive sufficient conditions on the pair,  $\mathbf{c}$  and  $\tau$ , so that (51) holds true.

Assume that the system is at the desired equilibrium, i.e, every Class- $j$  flow is sending its traffic over  $S_j$ . Let  $p_k^*$  be the equilibrium price (per unit rate) in  $S_k$ . The throughput received by a Class  $j$  flow (or anticipated by a Class  $j$  flow if it shifted to  $S_k$ ) is given by,

$$x_{jk}^* = \left( \frac{\tau_j}{p_k^*} \right)^{\frac{\beta}{\alpha}} \quad \text{and} \quad x_{Jk}^* = \left( \frac{T_l}{p_k^*} \right)^{\frac{\beta}{\alpha}}, \forall k. \quad (52)$$

The above results are due to the fact that the entry of a Class- $i$  flow into  $S_k$  may not significantly change its price,  $p_k^*$ , since there are large number of flows in  $S_k$ . In (52), the first result follows from Proposition 10, (34) and the assumption that  $T_j \in [T_s, T_l]$  when  $j < J$ , while the second one follows from Proposition 6. The link price  $p_k^*$  in  $S_k$ , follows from the above

results and the fact that rates of flows sharing a sub-network add up to its bandwidth allocation, is given by

$$p_k^* = \left( \frac{N_k}{c_k} \right)^{\frac{\beta}{\alpha}} \tau_k, \text{ if } k < J, \quad \text{and} \quad p_J^* = \left( \frac{N_J}{c_J} \right)^{\frac{\beta}{\alpha}} T_l \quad (53)$$

The payoff of Class  $j$  flow in  $S_k$ , from (8), is given by

$$\begin{aligned} F_{jk}(\mathbf{c}, \lambda) &= \frac{(x_{jk}^*)^{1-\alpha}}{1-\alpha} - \left( \frac{p_k^*}{\tau_j} \right)^\beta x_{jk}^* - \lambda_k, \\ &= A_{ik} \left( \frac{c_i}{N_i} \right)^{1-\alpha} - \lambda_i, \quad \forall k, \end{aligned} \quad (54)$$

where  $A_{ik} = \frac{\alpha}{1-\alpha} \left( \frac{\tau_i}{\tau_k} \right)^{\frac{\beta}{\alpha} (1-\alpha)}$  for  $i, k < J$ ,  $A_{iJ} = \frac{\alpha}{1-\alpha} \left( \frac{\tau_i}{T_l} \right)^{\frac{\beta}{\alpha} (1-\alpha)}, \forall i$ ,  $A_{Jk} = \frac{1}{1-\alpha} \left( \frac{T_l}{\tau_k} \right)^{\frac{\beta}{\alpha} (1-\alpha)}, k < J$  and  $A_{JJ} = \frac{1}{1-\alpha}$ . Also, (54) follows from (52) and (53).

The following lemma derives conditions on the pair  $(\mathbf{c}, \lambda)$  for (51) to hold true. Before stating the lemma, we introduce some notation. Let

$$l_{ik}(\mathbf{c}) = A_{ki} \left( \frac{c_i}{N_i} \right)^{1-\alpha} - A_{kk} \left( \frac{c_k}{N_k} \right)^{1-\alpha}. \quad (55)$$

$$u_{ik}(\mathbf{c}) = A_{ii} \left( \frac{c_i}{N_i} \right)^{1-\alpha} - A_{ik} \left( \frac{c_k}{N_k} \right)^{1-\alpha}, \quad (56)$$

**Lemma 14.** *Suppose the pair  $(\mathbf{c}, \lambda)$  satisfy the following conditions: if  $1 \leq k < J$ ,*

$$\frac{c_{k+1}}{c_k} \leq \frac{N_{k+1}}{N_k} \left( \frac{\tau_{k+1}}{\tau_k} \right)^{\frac{\beta}{\alpha}}, \quad (57)$$

$$\frac{c_J}{c_{J-1}} \leq \frac{N_J}{N_{J-1}} \left( \frac{\tau_l}{\tau_{J-1}} \right)^{\frac{\beta}{\alpha}}, \quad \sum_{j=1}^J c_j = c, \quad (58)$$

$$l_{k(k+1)}(\mathbf{c}) \leq \lambda_k - \lambda_{(k+1)} \leq u_{k(k+1)}(\mathbf{c}), \quad (59)$$

*Then, (51) hold true and the state where all the Class- $j$  flows choosing  $S_j, \forall j$ , is a Nash equilibrium.*

*Proof:* See Appendix ■

The system-value is sum of payoffs of all the flows, which is given by,

$$V_T(\mathbf{c}, \lambda) = \sum_i^J N_i F_{ii} = \sum_{i=0}^J N_i \left( A_{ii} \left( \frac{c_i}{N_i} \right)^{1-\alpha} - \lambda_i \right). \quad (60)$$

We must choose  $\mathbf{c}$  and  $\lambda$  that maximize (60) satisfying the NE conditions, (57)-(59). Let  $(\hat{\mathbf{c}}, \hat{\lambda})$  be one such optimal pair. Note that (60) is a decreasing function of toll vector,  $\lambda$ . Hence, from (58) and (64), we get

$$\hat{\lambda}_J = 0, \quad \text{and} \quad \hat{\lambda}_k = \sum_{i=k}^J l_{i(i+1)}. \quad (61)$$

Substituting the optimal toll values in (60), we get

$$\begin{aligned} V_T(\mathbf{c}) &= \frac{\bar{N}_J}{1-\alpha} \left( \frac{c_J}{\bar{N}_J} \right)^{1-\alpha} + \\ &\frac{\bar{N}_{J-1}\alpha}{1-\alpha} \left( \frac{c_{J-1}}{\bar{N}_{J-1}} \right)^{1-\alpha} \left( 1 - \frac{1}{\alpha} \left( \frac{T_l}{\tau_{J-1}} \right)^{\frac{\beta}{\alpha} (1-\alpha)} \right) + \\ &\sum_{k=1}^{J-2} \frac{\alpha \bar{N}_k}{1-\alpha} \left( \frac{c_k}{\bar{N}_k} \right)^{1-\alpha} \left( 1 - \left( \frac{\tau_{k+1}}{\tau_k} \right)^{\frac{\beta}{\alpha} (1-\alpha)} \right), \end{aligned} \quad (62)$$

where  $\bar{N}_k = \sum_{i=1}^k N_i$ . Then, define,

$$V_T = \max_{\mathbf{c}} V_T(\mathbf{c}) \quad \text{subject to} \quad (57) - (58). \quad (63)$$

We refer to  $V_T$  as *System value with tolling*. Now, we have the following proposition, which asserts that the system value achieved by the tolled multi-tier regime is superior to that of the untolled single tier regime.

**Proposition 15.** *The system value with tolling is no less than the value of single tier network game. i.e.,  $V_T \geq V_G$ . Also, the strict inequality holds if there exists a  $k < J$  such that*

$$\left(\frac{\bar{N}_J \bar{N}_k}{\bar{N}_J \bar{N}_k}\right)^{\frac{1}{\alpha}} \leq \left(\frac{T_l}{\tau_k}\right)^{\frac{\beta}{\alpha}} \left(1 - \left(\frac{\tau_{k+1}}{\tau_k}\right)^{\frac{\beta}{\alpha}(1-\alpha)}\right), \quad (64)$$

*Proof:* Suppose  $\mathbf{c}$  attains equality in (57)-(58), i.e a corner point of the constraint set. Note that the elements of  $\mathbf{c}$ , the bandwidths allocated to each subnetwork, that means to each flow class, is equal to the total bandwidth received by the corresponding flow class at the NE of the un-tolled single network game. Also, from (61) and (55), the optimal entrance toll in each subnetwork drops to zero. Then,  $V_T(\mathbf{c}) = V_G$ . Hence, we conclude that  $V_T \geq V_G$ .

Note that  $V_T(\mathbf{c})$  is strictly concave and hence, (63) has a unique maximizer. When (64) holds true, the unique maximizer lies in the interior of the constraint set of (63). Then,  $V_T > V_G$  which completes the proof. ■

Next, we derive a bound on the efficiency of the multi-tier tolling scheme. Let

$$\bar{\eta} = 1 + \alpha \sum_{k=1}^{J-1} \sum_{i=1}^k n_i. \quad (65)$$

where  $n_i = \frac{N_i}{N}$ . Then, we claim that

$$\eta_T = \frac{V_T}{V_S} \leq \min\{\eta_G, \bar{\eta}\}. \quad (66)$$

where  $\eta_G$  is the efficiency of single tier scheme without tolling. The claim can be proved as follows: Let  $\bar{c}_j = N_j \frac{c}{N}$  for all  $1 \leq j \leq J$ . Then,  $\bar{\mathbf{c}} = [\bar{c}_1, \dots, \bar{c}_J]$  lies in the feasible set of the optimization problem, (63). Then,  $V_T(\bar{\mathbf{c}}) \leq V_T$ . It can be shown that  $\frac{V_T(\bar{\mathbf{c}})}{V_S} < \bar{\eta}$  where  $V_S$  is given by (48). Therefore,  $\eta_T < \bar{\eta}$ . Also, from Proposition 15, we get that  $\eta_T \leq \eta_G$ . Together, we get the claim.

Note that,  $\bar{\eta}$ , does not depend on the ratio,  $\frac{T_s}{T_l}$ ; but it scales up with the number of classes in the system. Nevertheless,  $\eta_T$  is no more than the efficiency of the single tier networks without tolling. Therefore, we conclude that when the number of classes in the system is not arbitrarily large, the efficiency of multi-tier tolling schemes are superior to the single tier networks and, it does not scale up with the ratio,  $\frac{T_s}{T_l}$ . Note that there might be Nash equilibria other than the one stated by Lemma 14. Therefore, (66) may be better than the efficiency of the worst Nash equilibrium. Now, we present a numerical example to validate our analytical observations.

**Example-6:** Let two flow classes, namely Class 1 and Class 2, with disutility thresholds  $\tau_1 = T_s$  and  $\tau_2 = T_l$  are sharing a link with capacity  $c$  units. The link is partitioned into two subnetworks, namely  $S_1$  and  $S_2$ . Let  $N_i$  be the number of flows in Class  $i$  and define  $n_i = N_i/(N_1 + N_2)$ , for  $i = 1, 2$ . The optimal bandwidth allocation to subnetwork  $S_1$ , that

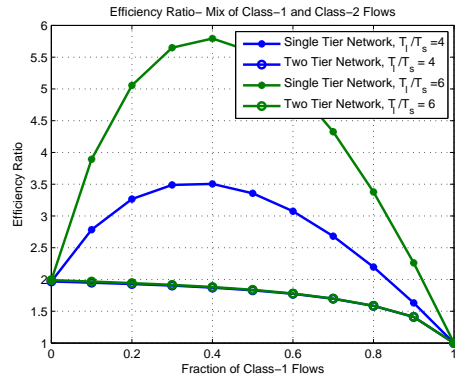


Fig. 7. Comparison of Efficiency Ratio ( $\eta$ ) between PMP scheme and Game in a network with price-insensitive flows and delay sensitive flows. Since  $V_S$  and  $V_G$  were negative in this example, a higher ratio is worse.

maximizes the system value with tolling, is given by

$$\hat{c}_1 = \frac{c}{1 + \frac{n_2}{n_1 n_2^\alpha} \left(1 - \left(\frac{T_l}{T_s}\right)^{\frac{\beta}{\alpha}(1-\alpha)}\right)^{-\frac{1}{\alpha}}} \vee \frac{c}{1 + \frac{n_2}{n_1} \left(\frac{T_l}{T_s}\right)^{\frac{\beta}{\alpha}}}.$$

Also, the optimal toll in  $S_1$  is given by  $\hat{\lambda}_1 = \left[\left(\frac{N_2}{c - \hat{c}_1}\right)^{\alpha-1} - \left(\frac{N_1}{\hat{c}_1} \left(\frac{T_s}{T_l}\right)^{\frac{\beta}{\alpha}}\right)^{\alpha-1}\right] \frac{\alpha}{\alpha-1}$ . Note that  $S_2$  has no entrance toll and the optimal allocation to  $S_2$  is  $\hat{c}_2 = c - \hat{c}_1$ . We define Efficiency Ratio ( $\eta_T$ ) here as the ratio of System-Value with tolling ( $V_T$ ) to Social optimum ( $V_S$ ). From (60) and  $V_S$ , (from (48)), we can show that

$$\eta_T = \frac{V_T}{V_S} = \frac{\alpha \left( (n_1 + n_2) \left(\frac{\hat{c}_1}{cn_1}\right)^{1-\alpha} + n_1 \left(\frac{\hat{c}_1}{cn_1}\right)^{1-\alpha} K \right)}{\left(1 + (\alpha - 1)(n_1 + n_2) \left(\frac{T_s}{T_l}\right)^\beta\right)},$$

where  $K = \left(1 - \left(\frac{T_l}{T_s}\right)^{\frac{\beta}{\alpha}(1-\alpha)}\right)$ .

In Figure (7), we have compared  $\eta$  attained using the PMP scheme versus that of a single-tier. We have used  $\alpha = 2$ ,  $\beta = 3$  and  $\left(\frac{T_l}{T_s}\right) = 4$  in our simulation. We observe that in spite of tolling, the PMP scheme always performs better than the single-tier scheme. Also, note that, unlike the single tier scheme, the efficiency of the PMP scheme does not scale with  $\frac{T_l}{T_s}$ .

## IX. CONCLUSION

In this paper we examined the consequences of the idea that a protocol is simply a way of interpreting Lagrange multipliers. We showed that flows could choose the interpretations, based on criteria such as delay or loss sensitivity. We determined the socially optimal protocol, as well as the choice that would result by flows taking their own selfish decisions. We showed that the social good is maximized by using the strictest possible price interpretation. However, based on different mixes of flow types a mix of interpretations could be the Nash equilibrium state. We characterized the loss of efficiency for some specific cases, and showed that a multi-tier network with tolling is capable of achieving superior system value. The result suggests the consideration of multiple tolled virtual networks, each geared towards a particular kind of flow. In the future we propose to explore the idea of virtual, tolled subnetworks further.

## X. APPENDIX-1

**Proof of Lemma 13:** The Nash equilibrium conditions, (51), are equivalent to

$$l_{ik}(\mathbf{c}) \leq \lambda_i - \lambda_k \leq u_{ik}(\mathbf{c}), \quad k > i, \forall i, \quad (67)$$

which follows from the definition of  $F_{ik}$  given by (54). Recall the definitions of,  $l_{ik}$  and  $u_{ik}$  from (55) and (56) respectively. Therefore, we prove the lemma by showing that (67) hold true when (57)-(59) are satisfied.

Suppose (57)-(59) are true. Then, it is easy to observe that  $l_{ik} \leq u_{ik}, \forall k > i$ . Also, we have

$$\sum_{t=k}^{m-1} l_{t(t+1)} \leq \lambda_k - \lambda_m, \forall m > k, \forall k. \quad (68)$$

From the definitions of  $l_{ik}$ 's and the fact that  $\tau_i < \tau_k$  if  $i < k$ , it is easy to show that

$$l_{k(k+j)} - l_{k(k+j-1)} \leq l_{(k+j-1)(k+j)}, \quad (69)$$

for  $k < J$  and  $1 < j \leq J - k$ . Then, we have,

$$\begin{aligned} l_{km} &= l_{k(k+1)} + (l_{k(k+2)} - l_{k(k+1)}) + \dots + (l_{km} - l_{k(m-1)}) \\ &\leq \sum_{t=k}^{m-1} l_{t(t+1)} \leq \lambda_k - \lambda_m. \end{aligned} \quad (70)$$

In similar fashion, we can show that  $u_{km} \geq \lambda_k - \lambda_m$ . Then, (67) is proved and hence the lemma.

## REFERENCES

- [1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.
- [2] S. H. Low and D. E. Lapsley, "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, pp. 861–875, December 1999.
- [3] S. Kunniyur and R. Srikant, "A time-scale decomposition approach to adaptive ECN marking," in *Proceedings of IEEE Infocom*, Anchorage, AK, April 2001.
- [4] R. Srikant, *The Mathematics of Internet Congestion Control*. Birkhauser, 2004.
- [5] S. Shakkottai and R. Srikant, "Network optimization and control," *Found. Trends Netw.*, vol. 2, no. 3, pp. 271–379.
- [6] R. Stevens, *TCP/IP Illustrated, Volume 1*. Addison-Wesley, 1994.
- [7] T. Kelly, "Scalable TCP: Improving performance in highspeed wide area networks," December 2002.
- [8] D. X. Wei, C. Jin, S. H. Low, and S. Hegde, "FAST TCP: motivation, architecture, algorithms, performance," *IEEE/ACM Transactions on Networking*, Dec. 2006.
- [9] S. Liu, T. Başar, and R. Srikant, "TCP-Illinois: a loss and delay-based congestion control algorithm for high-speed networks," in *valuetools '06: Proceedings of the 1st international conference on Performance evaluation methodologies and tools*, Pisa, Italy, October 2006.
- [10] A. Tang, J. Wang, S. Hegde, and S. Low, "Equilibrium and fairness of networks shared by TCP Reno and Vegas/FAST," *Telecommunication Systems*, vol. 30, no. 4, pp. 417–439, 2005.
- [11] A. Tang, J. Wang, S. Low, and M. Chiang, "Equilibrium of heterogeneous congestion control: Existence and uniqueness," *IEEE/ACM Transactions on Networking*, vol. 15, no. 4, pp. 824–837, August 2007.
- [12] A. Tang, X. Wei, S. Low, and M. Chiang, "Equilibrium of heterogeneous congestion control: Optimality and stability," *IEEE/ACM Transactions on Networking*, vol. 18, no. 3, pp. 844–857, June 2010.
- [13] J. Mo, R. La, V. Anantharam, and J. Walrand, "Analysis and comparison of TCP Reno and Vegas," in *IEEE Infocomm*, 1999, pp. 1556–1563.
- [14] E. Altman, R. El-Azouzi, Y. Hayel, and H. Tembine, "The evolution of transport protocols: An evolutionary game perspective," *Computer Networks*, vol. 53, no. 10, pp. 1751–1759, 2009.

- [15] M. Podlesny and S. Gorinsky, "RD network services: differentiation through performance incentives," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, 2008, pp. 255–266.
- [16] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, October 2000.
- [17] A. M. Odlyzko, "Paris Metro Pricing for the Internet," in *Proceedings of the ACM Conference on Electronic Commerce (EC'99)*, 1999.
- [18] S. Shakkottai, R. Srikant, A. Ozdaglar, and D. Acemoglu, "The price of simplicity," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1269–1276, September 2008.
- [19] C.-K. Chau, Q. Wang, and D.-M. Chiu, "On the viability of Paris Metro Pricing for communication and service networks," in *Proceedings of IEEE Infocom*, San Diego, CA, March 2010.



**Vinod Ramaswamy** received his Bachelor of Engineering degree in electronics and communication from Kerala University, India, in 2004 and his M.Tech. degree in electrical engineering, with specialization in communication systems, from Indian Institute of Technology, Madras, India, in 2006. He is currently working towards the Ph.D. degree in Electrical Engineering at Texas A&M University, College Station.

His research interests include game theory, optimization methods, and their application to communication networks.



**Diganto Choudary** received his Bachelors of Science degree in Electrical Engineering with a minor in Mathematics from Texas A&M University in 2012. He has received the distinction of Undergraduate Research Fellow and Engineering Scholar.

He is currently working in Maxim Integrated in Dallas TX. This work was done as part of his Undergraduate Summer Research Program. His research interests include peer-to-peer systems, queuing theory, connection between information theory and electromagnetic theory, and congestion control.



**Srinivas Shakkottai** (S '00-M '08) received his Bachelor of Engineering degree in electronics and communication engineering from the Bangalore University, India, in 2001 and his M.S. and Ph.D degrees from the University of Illinois at Urbana-Champaign in 2003 and 2007, respectively, both in electrical engineering. He was Postdoctoral Scholar at Stanford University until December 2007, and is currently an Assistant Professor at the Dept. of ECE, Texas A&M University. He has received the Defense Threat reduction Agency Young Investigator Award (2009) and the NSF Career Award (2012) as well as research awards from Cisco (2008) and Google (2010).

His research interests include content distribution systems, pricing approaches to network resource allocation, game theory, congestion control, and the measurement and analysis of Internet data.