Congestion and Its Role in Network Equilibrium

Richard T. B. Ma and Vishal Misra

***Abstract*—In this paper, we develop the notion of *congestion equilibrium* in large scale networks, with the specific goal of understanding the modern multiparty Internet ecosystem comprising of content providers, ISPs and users. We show that the concept of "congestion-taking" is analogous to the concept of "price-taking" in classical market economics. With a wide variety of congestion metrics and under very mild assumptions on the congestion dynamics, we characterize various properties of congestion equilibria and develop algorithms to compute them for large scale networks. Our work provides a new way to model and analyze modern large scale network-economic systems that have a complex interaction of engineering and economics.**

***Index Terms*—Congestion Equilibrium, Internet Economics.**

I. Introduction

The INTERNET has been growing into a network of thousands of interconnected ISPs that autonomously and strategically interact with each other. To model the evolution of the Internet, equilibrium concepts have been extensively used to capture the system characteristics to a first approximation. In particular, Nash equilibrium [19] is the dominating solution concept that models the detailed strategic interactions. However, applying this delicate solution concept is often impractical in the context of large scale networks like the Internet. First of all, the "common knowledge" [21] assumption behind Nash equilibrium is not valid in reality, because ISPs normally do not disclose private information, e.g., available resources and routing strategies. Even with the "common knowledge", the strategy space increases exponentially with the scale of the system, and solving Nash equilibrium has been shown to be computationally expensive [7], [11] even under two players. In contrast to Nash equilibrium, competitive equilibrium in an exchange economy [16] is a macroscopic solution that models many small competing players in a perfect market. Under a competitive equilibrium, any individual player's consumption or production decision of a commodity does not move the equilibrium/market price of that commodity and each player determines the optimal strategy solely based on this market price. Thus, the market price plays the role of "common knowledge" and under a minor assumption of "price-taking", each player can optimize its strategy in a decentralized manner. In practice, competitive equilibrium gives more plausible results for perfect competition, i.e., no single player has substantial market power, while still captures the network effect (or externality) among the players via price.

Manuscript received 15 December 2011; revised 20 May 2012.

R. Ma is with Advanced Digital Sciences Center, Illinois at Singapore (e-mail: tbma@adsc.com.sg). He is also with School of Computing, National University of Singapore (e-mail: tbma@comp.nus.edu.sg).

V. Misra is with Department of Computer Science, Columbia University (e-mail: misra@cs.columbia.edu).

Digital Object Identifier 10.1109/JSAC.2012.121210.

Unlike an exchange economy, modern networking systems do not have commodities to exchange. Instead, by utilizing packet switching and statistical multiplexing, they share resources, e.g., bandwidth capacity, among users. In this case, users affect each other's performance via network congestion, a type of negative network externality in the system. Analogous to the market prices of commodities seen in competitive markets, network players often induce and see a common level of congestion. Inspired by the concept of competitive equilibrium and its "pricing-taking" assumption, we introduce the notation of *congestion equilibrium* for network systems and the corresponding "congestion-taking" assumption. Congestion equilibrium is suitable to model large scale networks where players' collective decisions, e.g., routing decisions, determine the level of congestion in different parts of the network, while any individual player's decision does not affect the level of congestion much.

To illustrate this concept, we develop a framework of the Internet ecosystem under which content providers (CPs) compete for a last-mile bottleneck capacity of an ISP so as to service their users. The ISP implements a Paris Metro Pricing (PMP) [20] type of service differentiation by dividing its capacity into a premium and an ordinary class, and CPs get charged extra for sending traffic in the premium class. For any single service class, we derive a user-level congestion equilibrium, under which users compete for the capacity of the service class and reach a level of congestion in equilibrium. By using the user-level equilibrium as a building block, we further derive a CP-level congestion equilibrium across multiple service classes, where CPs compete with other CPs in the same service class and strategically choose the service class to maximize individual utilities. Our major findings and contributions include:

- "Congestion-taking" needs to be assumed for the players under congestion equilibria (Assumption 1 and 5).
- A user-level congestion equilibrium can be uniquely characterized by the user demand for different content and the congestion dynamics of the system (Theorem 1).
- The user-level congestion equilibrium has monotonicity and scaling properties (Theorem 2 and 3).
- Our model is applicable to various congestion metrics, including the $M/G/1$ delay and the throughput resulted from rate or/and congestion control mechanisms.
- Under a PMP-type of service differentiation provided by an ISP, we formulate a strategic game for the CPs to choose its preferred service class and derive the corresponding CP-level congestion equilibrium (Section IV).
- We characterize the uniqueness of the CP-level congestion equilibrium under a special case (Theorem 6) and

propose algorithms to efficiently calculate the CP-level congestion equilibria for general cases.

Our new notion of congestion equilibrium sheds new light on modeling and analyzing practical and plausible network equilibrium for large scale networking systems, which avoids the complexity and impracticality of Nash equilibrium.

II. Related Work

The concept of congestion happens naturally in transportation networks. Wardrop [25] pointed out that congestion can happen if users strategically choose individual routes to optimize their own utilities. In particular, Wardrop equilibrium is a type of Nash equilibrium, which may lead to an inefficient system state. An example is the Braess's Paradox [5] where adding extra capacity to a network can in some cases reduce overall performance. Rosenthal [22] first proposed a framework of congestion games, which is a special case of potential games [18] where the existence of a pure-strategy Nash equilibrium is guaranteed.

Our model of the Internet ecosystem and network congestion has two major differences from classic congestion models of transportation networks and congestion games. First, transportation networks often involve two parties: individual drivers and route/capacity planners. However, the Internet naturally involves three parties, i.e., the ISPs, CPs and users, which forms a two-sided market [3]. The dynamics of the Internet and the corresponding player strategies are quite different from the transportation networks. The cause of congestion in the Internet depends on many factors: under a user-level equilibrium, the congestion depends on users' demand for different content and the physical congestion dynamics of the network system; under a CP-level equilibrium, the congestion also depends on ISPs' service differentiations and the CPs' decisions for joining various service classes. Second, instead of using Nash equilibrium as the solution concept for the system, we use the concept of congestion equilibrium that is based on a "congestion-taking" assumption of the players.

The "congestion-taking" assumption is similar to the "pricing-taking" assumption under well-studied exchange economies [16] in microeconomics. In the networking area, Kelly [13] [12] studied resource allocation and pricing problems and showed that if resources are shared proportionally and under a pricing-taking assumption, the resulting equilibrium will be optimal, i.e. the social welfare is maximized.

The computational complexity of the Nash equilibrium has recently been shown to be PPAD-complete [7] even under two players and for approximation [11]. Many refined solution concept, such as subgame perfect equilibrium [21] and trembling hand perfect equilibrium [23], have been proposed to capture more nuanced characteristics that are not addressed in Nash equilibrium. However, since the assumption of "common knowledge" is often invalid for the players in large scale networks, we make a more realistic "congestion taking" assumption of the players and use a "rougher" solution concept than Nash equilibrium. Congestion equilibrium also has some similarity with the concept of *correlated equilibrium* proposed by Robert Aumman [4], which is a superset of Nash equilibrium. In a correlated equilibrium, each player chooses her action according to her observation of the value of a same public signal. Under our context, the public signal is a metric of congestion resulted from the physical dynamics of a network system. The difference is that, in a correlated equilibrium, "common knowledge" is still assumed and each player's observation can be subjective and different from each other. The observation of congestion in the system, however, is objective and has the same value for all the players.

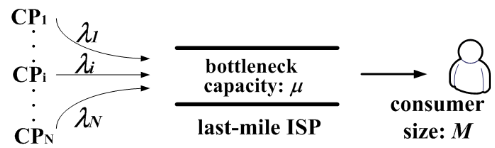


Fig. 1. Contention at the last-mile bottleneck link.

In one recent related work, we applied the idea of congestion equilibrium in modeling ISP competition in monopolistic and oligopolistic markets and inform the desirability of network neutrality regulation [15]. Notice that, our notation of congestion equilibrium and the theoretical results are more fundamental than what we derived in [15], which only focused on throughput as the congestion metric and the corresponding system mechanism for throughput allocation.

III. User-level Congestion Equilibrium

We consider a model of the Internet ecosystem with three parties: 1) CPs, 2) a last-mile ISP and 3) end users. We focus on a fixed user group in a targeted geographic region. We denote M as the number of users in the region[1]. Each user subscribes to an Internet access service via an ISP. We consider the scenario of a single last-mile ISP I that provides the Internet access for the users. As we will show later, our model can be easily extended to model a group $\mathcal{I}$ of last-mile ISPs. We denote $\mathcal{N}$ as the set of CPs from which the users request content. We define $N = |\mathcal{N}|$ as the number of CPs. Our model does not include the backbone ISPs for two reasons. First, the bottleneck of the Internet is often at the last-mile connection towards the users [10], both wired and wireless. We focus on the access or so-called eyeball ISPs [14] that provide the bottleneck last-mile towards the users. Second, the recent concern on network neutrality manifests itself in the cases where the last-mile ISPs intended to differentiate services and charge CPs, e.g. Apple and Google, for service fees [6]. Our model captures the non-neutral pricing and service differentiation behaviors of the last-mile ISPs, if network neutrality turns out to be unnecessary and is not imposed.

We denote μ as the last-mile bottleneck capacity towards the users in the region. Figure 1 depicts the contention at the bottleneck among different flows from the CPs. We denote λ_i as the aggregate throughput rate from CP i to the users. For any set $\mathcal{N}$ of CPs, we define $\Lambda = (\lambda_1, \cdots, \lambda_N)$ and the total throughput rate as $\lambda_{\mathcal{N}} = \sum_{i\in\mathcal{N}} \lambda_i$. Because users initiate downloads and retrieve content from the CPs, we first model

[1]Note that M can also be interpreted as the *average* or *peak* number of users accessing the Internet simultaneously in the region, which will scale with the total number of actual users in a region. This does not change the nature of any of the results we describe subsequently, but gives a more realistic interpretation of the congestion and throughput rates in equilibrium.

the user demands so as to characterize the CPs' throughput rates. If the ISP only provides a single service class with all its capacity μ, given a set $\mathcal{N}$ of CPs and a group of M users, we denote the system as a triple $(M, \mu, \mathcal{N})$.

A. User Throughput Demand as a Function of Congestion

We denote $\hat{\theta}_i$ as the *unconstrained throughput* for a typical user of CP i. For instance, the unconstrained throughput for the highest quality Netflix streaming movie is about 5 Mbps [2], and given an average query page of 20 KB and an average query response time of .25 seconds [1], the unconstrained throughput for a Google search is about 600 Kbps, or just over $1/10^{th}$ of Netflix. We denote $\alpha_i \in (0,1]$ as the percentage of users that ever access CP i's content, which models the popularity of the content of CP i. We define $\hat{\lambda}_i = \alpha_i M \hat{\theta}_i$ as the unconstrained throughput of CP i. Without contention, CP i's throughput λ_i equals $\hat{\lambda}_i$. However, when the capacity μ cannot support the unconstrained throughput for all the CPs, e.g., $\mu < \sum_{i\in\mathcal{N}} \hat{\lambda}_i$, each CP's rate $\lambda_i \leq \hat{\lambda}_i$. In general, CP i's rate λ_i depends on the level of congestion in the network. We denote a non-negative real number $\phi \in \mathbb{R}_+$ as the level of congestion caused by the set of competing CPs $\mathcal{N}$. We will see that ϕ can represent various congestion metrics in general. Each CP's rate can be expressed as a function[2] of the user population M and the system congestion ϕ as

$$\lambda_i(M, \phi) = \alpha_i M \rho_i(\phi), \tag{1}$$

where $\rho_i(\phi)$ can be interpreted as the per-user achievable rate under congestion ϕ. We define the throughput vector as a function of M and ϕ as $\Lambda(M,\phi) = (\lambda_1(M,\phi), \cdots, \lambda_N(M,\phi))$.

Assumption *1 (Congestion Taking by Users):* For any CP i, $\rho_i : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous and non-increasing function that satisfies $\rho_i(0) = \hat{\theta}_i$ and $\lim_{\phi\to\infty} \rho_i(\phi) = 0$.

Although the aggregate user demand for certain content affects the system congestion as well as the throughput of other content, the above assumption implies that the aggregate user demand for any content only depends on their experienced level of congestion. Since each individual user's demand decision, whether to continue or stop downloading, really depends on the level of congestion the user experiences, it is reasonable to make the above congestion-taking assumption for the aggregate user demand for any type of content.

B. Congestion as a Function of Throughput and Capacity

When flows share the same bottleneck link, they compete for capacity and therefore, induce a level of congestion in the system. The system congestion can be viewed as the outcome of a continuous function Φ of the throughput Λ of the CPs and the system capacity μ. Without specifying the congestion metric, we make the following assumption on the congestion dynamics Φ induced by the underlying network system.

Assumption *2 (Congestion Monotonicity of Throughput):* $\Phi(\Lambda, \mu) : \mathbb{R}^N_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function of the throughput rates $\Lambda \in \mathbb{R}^N_+$. For any rates $\Lambda_1 \leq \Lambda_2$, it satisfies

$$\Phi(\Lambda_1, \mu) \leq \Phi(\Lambda_2, \mu), \quad \forall \mu \geq 0.$$

Moreover, for any subset $\mathcal{N}' \subset \mathcal{N}$, let Λ' be the projection of Λ in the subspace $\mathbb{R}^{N'}_+$, it also satisfies

$$\Phi(\Lambda', \mu) \leq \Phi(\Lambda, \mu), \quad \forall \mu \geq 0.$$

Assumption 2 states that under a fixed system capacity, the increase in congestion is caused by the increase of the number of CPs or/and the throughput from them.

Assumption *3 (Congestion Monotonicity of Capacity):* $\Phi(\Lambda, \mu) : \mathbb{R}^N_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function of the system capacity μ. For any capacity $\mu_1 \geq \mu_2$, it satisfies

$$\Phi(\Lambda, \mu_1) \leq \Phi(\Lambda, \mu_2), \quad \forall \Lambda \geq 0.$$

Assumption 3 states that under fixed throughput of the CPs, more congestion is caused by the decrease of system capacity.

Assumption *4 (Congestion Monotonicity of System Scale):* For any $\xi \geq 1$, the system congestion Φ satisfies

$$\Phi(\xi\Lambda, \xi\mu) \leq \Phi(\Lambda, \mu).$$

Assumption 4 states that if the user size and system capacity scale up linearly at the same rate, the level of system congestion would not become worse. It captures the efficiency of statistical multiplexing in the modern networking systems.

C. Properties of the User-Level Congestion Equilibrium

The demand function $\rho_i(\cdot)$ maps a congestion level ϕ to throughput λ_i; the system congestion dynamics $\Phi(\cdot,\cdot)$ maps any fixed demands Λ and capacity μ to a level of congestion. The interplay between the two determines the system congestion and throughput in a *user-level congestion equilibrium*.

Definition *1:* A congestion level φ is a *user-level congestion equilibrium* of the system $(M, \mu, \mathcal{N})$ if $\Phi(\Lambda(M, \varphi), \mu) = \varphi$.

Theorem *1 (Uniqueness of Congestion Equilibrium):* Under Assumption 1 and 2, a system $(M, \mu, \mathcal{N})$ has a unique user-level congestion equilibrium φ.

Remark: The convergence to a user-level congestion equilibrium happens naturally in practice. If the instantaneous congestion is higher than φ, some users will back off which leads to lower throughput demand; if the instantaneous congestion is lower than φ, better performance will attract more users to demand for throughput which increases congestion. This will become clear when we break down the function ρ_i and look into more detailed user demand in the next subsection.

By Theorem 1, we denote $\varphi = \varphi(M, \mu, \mathcal{N})$ as the unique user-level congestion equilibrium and define $\Lambda(M, \mu, \mathcal{N}) = \Lambda(M, \varphi)$ as the corresponding throughput of the CPs.

Theorem *2 (Monotonicity of Congestion Equilibrium):* Under Assumption 1 to 3, for any capacity $\mu_1 \geq \mu_2 \geq 0$, user population $0 \leq M_1 \leq M_2$ and $\mathcal{N}_1 \subseteq \mathcal{N}_2$, the user-level congestion equilibrium satisfies

$$\varphi(M_1, \mu_1, \mathcal{N}_1) \leq \varphi(M_2, \mu_2, \mathcal{N}_2).$$

Moreover, for any fixed population M, the throughput satisfies

$$\lambda_i(M, \mu_1, \mathcal{N}_1) \geq \lambda_i(M, \mu_2, \mathcal{N}_2), \quad \forall i \in \mathcal{N}_1.$$

Remark: Given fixed μ and $\mathcal{N}$, $\lambda_i(M, \mu, \mathcal{N})$ might not be monotonic in M. When M increases, although the per-capita

[2]We use λ_i as a fixed throughput rate and $\lambda_i(\cdot)$ as a function.

capacity μ/M decreases and the congestion φ increases, the throughput of congestion-insensitive CPs could increase.

Theorem *3 (Scaling of User-level Congestion Equilibrium):* Under Assumption 1, 2 and 4, for any $\xi \geq 1$, the user-level congestion equilibrium satisfies

$$\varphi(\xi M, \xi\mu, \mathcal{N}) \leq \varphi(M, \mu, \mathcal{N}), \quad \text{and}$$
$$\Lambda(\xi M, \xi\mu, \mathcal{N}) \geq \xi\Lambda(M, \mu, \mathcal{N}).$$

In particular, if $\Phi(\xi\Lambda, \xi\mu) = \Phi(\Lambda, \mu)$ in the condition of Assumption 4, the congestion and throughput are homogenous functions of degree 0 and -1 respectively, i.e., $\varphi(M, \mu, \mathcal{N}) = \varphi(\xi M, \xi\mu, \mathcal{N})$ and $\Lambda(M, \mu, \mathcal{N}) = \xi^{-1}\Lambda(\xi M, \xi\mu, \mathcal{N}), \forall \xi > 0$.

Remark: Under the scaling condition $\Phi(\xi\Lambda, \xi\mu) = \Phi(\Lambda, \mu)$ and by the monotonicity property of Theorem 2, one can also express the congestion equilibrium as a continuous non-increasing function of the per-capita capacity μ/M.

D. Illustrations of Congestion Metrics and Dynamics

In this subsection, we illustrate various congestion metrics that can be modeled by our user-level congestion equilibrium.

FIFO queueing delay: In the simplistic example, the ISP only performs a first-come-first-serve service discipline on the packet flows. Since all data packets encounter the same average queueing delay under the FIFO scheduler, we can naturally measure the level of congestion by the queueing delay of the system. Under an M/G/1 model, by the Pollaczek-Khinchine mean formula, the congestion function is

$$\Phi(\Lambda, \mu) = \frac{\lambda E[S^2]}{2(1-\rho)} = \frac{\lambda_{\mathcal{N}} E[S^2]}{2(1-\lambda_{\mathcal{N}}/\mu)},$$

where $E[S^2]$ is the second moment of packet service time. In particular, under an M/M/1 model, the congestion function is

$$\Phi(\Lambda, \mu) = \frac{\lambda_{\mathcal{N}}}{\mu(\mu - \lambda_{\mathcal{N}})}.$$

In these cases, the congestion reduces strictly when the user size and system capacity scale up linearly at the same rate.

User throughput under proportional rate control: Besides delay-based congestion metrics, the achievable throughput of active users is another measure of congestion. We denote $\theta_i(\phi)$ as the achievable throughput of an active user of CP i under congestion level ϕ and define $\Theta = (\theta_1, \cdots, \theta_N)$. The average per-user achievable rate $\rho_i(\phi)$ can be expressed as $\rho_i(\phi) = d_i(\phi)\theta_i(\phi)$, where $d_i(\phi)$ denotes the percentage of users still being active under congestion ϕ. Because given the CP aggregate throughput Λ, or ρ_i of all CP i, we can derive the user throughput Θ, we will see that some congestion metrics $\Phi(\Lambda, \mu)$ can be expressed as $\Phi(\Theta)$, a function of the user throughput. In general, $d_i(\cdot)$ and $\theta_i(\cdot)$ are non-increasing functions with $\theta_i(0) = \hat{\theta}_i$, $d_i(0) = 1$ and $\lim_{\phi\to\infty} \theta_i(\phi) = 0$, which implies that when least congested, i.e., $\phi = 0$, each active user should receive the unconstrained throughput rate $\hat{\theta}_i$ and 100% of the interested users remain active in the system. Later, we will show that $d_i(\cdot)$ also capture the users' demand sensitivity to the system congestion.

The equilibrium user throughput depends on the rate control mechanism of the network. A rate control mechanism can be a flow control mechanism, e.g. CBR and VBR, under which the bottleneck link decides the rates for each flow in a centralized manner, or a window-based end-to-end congestion control mechanism, e.g. TCP, under which each flow maintains a sliding window and adapts its size based on implicit feedback from the network, e.g. round-trip time.

Suppose the bottleneck ISP perform a rate control mechanism that assign throughput for flows proportional their maximum demand, i.e., $\theta_i : \theta_j = \hat{\theta}_i : \hat{\theta}_j$ for all $i, j \in \mathcal{N}$, a system-wide congestion metric can be defined as

$$\Phi(\Lambda, \mu) = \Phi(\Theta) = \frac{\hat{\theta}_i}{\theta_i} - 1, \quad \forall i \in \mathcal{N}. \tag{2}$$

User throughput under end-to-end congestion control: Due to the end-to-end design principle of the Internet [9], congestion control has been implemented by window-based protocols, i.e. TCP and its variations. Mo and Walrand [17] showed that a class of α-proportional fair solutions can be implemented by window-based end-to-end protocols. Among the class of α-proportional fair solutions, the max-min fair allocation, a special case with $\alpha = \infty$, is the result of the AIMD mechanism of TCP [8]. Differing round trip times, receiver window sizes and loss rates can result in different bandwidths, but to a first approximation, TCP provides a max-min fair allocation of available bandwidth amongst flows.

Under an end-to-end congestion control mechanism that achieves the max-min fair allocations, a system-wide congestion metric could be defined as

$$\Phi(\Lambda, \mu) = \Phi(\Theta) = \frac{1}{\max\{\theta_i : i \in \mathcal{N}\}}.$$

Notice that for the proportional rate control and the end-to-end congestion control mechanisms, one can verify that the congestion metrics satisfy $\Phi(\xi\Lambda, \xi\mu) = \Phi(\Lambda, \mu)$ for all $\xi > 0$.

IV. MULTI-CLASS CP-LEVEL CONGESTION EQUILIBRIUM

In this section, we consider an eyeball ISP with capacity μ, which can be a retail residential ISP, e.g., Comcast and Time Warner Cable, or a mobile operator, e.g., Verizon and AT&T. Regardless of being a wired or wireless provider, it serves as the last-mile service provider for the users. We assume that the ISP is allowed to allocate a fraction $\kappa \in [0, 1]$ of its capacity to serve premium CPs and charge them at a rate $c \in [0, \infty)$ (dollar per unit traffic). For a wired ISP, κ can be interpreted as the percentage of capacity deployed for private peering points that charge a fee of c per unit incoming traffic and $1 - \kappa$ can be interpreted as the percentage of capacity deployed for public peering points where incoming traffic is charge-free. For a wireless ISP, κ can be interpreted as the percentage of capacity devoted for the premium traffic that will be charged at a rate of c. The pair of parameters (κ, c) can also be thought of a type of Paris Metro Pricing (PMP) [20], [24], where one ordinary and another premium service class have capacities of $(1 - \kappa)\mu$ and $\kappa\mu$ and charge 0 and c respectively. In reality, content might be delegated via content distribution networks (CDNs), e.g. Akamai, or backbone ISPs, e.g. Level3 is a major tier-1 ISP that delivers Netflix traffic towards regional ISPs. Therefore, in practice, the charge c might be imposed on the delivering ISP, e.g. Level3, and then be recouped from the CP, e.g. Netflix, by its delivering ISP, e.g. Level3. Our model does not assume any form of the implementation.

We denote $\mathcal{O}$ and $\mathcal{P}$ as the two disjoint sets of CPs that join the ordinary and premium class respectively. As a consequence of service differentiation, virtually, the original system $(M, \mu, \mathcal{N})$ breaks into two independent subsystems $(M, (1-\kappa)\mu, \mathcal{O})$ and $(M, \kappa\mu, \mathcal{P})$. We denote $\varphi_{\mathcal{O}} = \varphi(M, (1-\kappa)\mu, \mathcal{O})$ and $\varphi_{\mathcal{P}} = \varphi(M, \kappa\mu, \mathcal{P})$ as the congestion level in both service classes under a user-level congestion equilibrium respectively. We denote v_i as CP i's per unit traffic revenue. This revenue can be generated by advertising for media clients, e.g Google, or by selling products to online users, e.g. Amazon, or by providing services to users, e.g. Netflix and e-banking. Our model does not assume how the revenue is generated either. Each CP i's utility function u_i can be expressed as

$$u_i(\lambda_i) = \begin{cases} v_i \lambda_i(M, \varphi_{\mathcal{O}}) & \text{if } i \in \mathcal{O}, \\ (v_i - c)\lambda_i(M, \varphi_{\mathcal{P}}) & \text{if } i \in \mathcal{P}. \end{cases} \tag{3}$$

We assume that a CP chooses to join the premium service class only if it provides higher utility than the ordinary service class.

A. Content Provider's Best Response

Given the ISP's decision κ and c, each CP chooses whether to join the ordinary service class $\mathcal{O}$ or the premium class $\mathcal{P}$.

Lemma *1:* Given a fixed set $\mathcal{O}$ of CPs in the ordinary class and a fixed set $\mathcal{P}$ of CPs in the premium class, a new CP i's optimal strategy is to join the premium service class, if

$$(v_i - c)\ \rho_i\big(\varphi_{\mathcal{P}\cup\{i\}}\big) > v_i\ \rho_i\big(\varphi_{\mathcal{O}\cup\{i\}}\big). \tag{4}$$

Lemma 1 states that a CP will join the premium service class if that results higher profit, which is per-unit flow profit ($v_i - c$ for the premium class) multiplied by the per capita throughput ρ_i. Since $\rho_i(\cdot)$ is non-negative, Lemma 1 also implies that, if $v_i \le c$, then CP i's optimal strategy is always to join the ordinary service class. The above decision is clear for a CP only if all other CPs have already made their choices. To treat all CPs equally, we model the decisions of all CPs as a simultaneous-move game as part of a two-stage game.

B. Two-Stage Strategic Game and Nash Equilibrium

We model the strategic behavior of the ISP and the CPs as a two-stage game, denoted as a quadruple $(M, \mu, \mathcal{N}, I)$.

1) *Players*: The ISP I and the set of CPs $\mathcal{N}$.
2) *Strategies*: ISP I chooses a strategy $s_I = (\kappa, c)$. Each CP i chooses a binary strategy of whether to join the premium class. The CPs' strategy profile can be written as $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$, where $\mathcal{O} \cup \mathcal{P} = \mathcal{N}$ and $\mathcal{O} \cap \mathcal{P} = \emptyset$.
3) *Rules*: In the first stage, ISP I decides $s_I = (\kappa, c)$ and announces it to all the CPs. In the second stage, all the CPs make their binary decisions simultaneously and reach a joint decision $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$.
4) *Outcome*: The set $\mathcal{P}$ of the CPs shares a capacity of $\kappa\mu$ and the set $\mathcal{O}$ of the CPs shares a capacity of $(1-\kappa)\mu$. Each CP $i \in \mathcal{O}$ gets a rate $\lambda_i(M, (1-\kappa)\mu, \mathcal{O})$ and each CP $j \in \mathcal{P}$ gets a rate $\lambda_j(M, \kappa\mu, \mathcal{P})$.
5) *Payoffs*: Each CP i's payoff is defined by the utility $u_i(\lambda_i)$ in Equation (3). The ISP's payoff is the revenue $c\lambda_{\mathcal{P}}$ received from the premium class.

If we regard the set of CPs as a single player that chooses a strategy $s_{\mathcal{N}}$, our two-stage game can be thought of a Stackelberg game [21]. In this game, the first-mover ISP can take all the best-responses of the CPs into consideration and derive its optimal strategy s_I using *backward induction* [16]. Given any fixed strategy $s_I = (\kappa, c)$, the CPs derive their best strategies under a simultaneous-move game, denoted as $(M, \mu, \mathcal{N}, s_I)$. We denote $s_{\mathcal{N}}(M, \mu, \mathcal{N}, s_I) = (\mathcal{O}, \mathcal{P})$ as a strategy profile of the CPs under the game $(M, \mu, \mathcal{N}, s_I)$. Technically speaking, when $\kappa = 0$ or 1, there is only one service class. When $\kappa = 0$, we define the trivial strategy profile as $s_{\mathcal{N}} = (\mathcal{N}, \emptyset)$; when $\kappa = 1$, although there is not a physical ordinary class, we define the trivial strategy profile as $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{N} \backslash \mathcal{O})$, with $\mathcal{O} = \{i : v_i \le c, i \in \mathcal{N}\}$ which defines the set of ISPs that cannot afford to join the premium class. Based on Lemma 1, we can define an equilibrium in the sense of a Nash or a CP-level congestion equilibrium.

Definition *2:* A strategy profile $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$ is a Nash equilibrium of a game $(M, \mu, \mathcal{N}, s_I)$, if

$$\begin{cases} (v_i - c)\ \rho_i\big(\varphi_{\mathcal{P}\cup\{i\}}\big) \le v_i\ \rho_i\big(\varphi_{\mathcal{O}}\big), & \forall i \in \mathcal{O}, \\ (v_i - c)\ \rho_i\big(\varphi_{\mathcal{P}}\big) > v_i\ \rho_i\big(\varphi_{\mathcal{O}\cup\{i\}}\big), & \forall i \in \mathcal{P}. \end{cases} \tag{5}$$

C. CP-Level Congestion Equilibrium

Notice that a CP's joining decision to a service class might increase the congestion level and reduce the throughput of flows of that service class; however, if the number of CPs in a service class is big and no single CP's traffic will dominate, an additional CP i's effect will be marginal. Analogous to the congestion-taking (Assumption 1) at the user-level, we make the following congestion-taking assumption for the CPs.

Assumption *5 (Congestion Taking by the CPs):* For any service class $\mathcal{X}$, any CP $i \notin \mathcal{X}$ uses $\varphi_{\mathcal{X}}$ as an estimate of the ex-post congestion $\varphi_{\mathcal{X}\cup\{i\}}$ it will face in the decision-making. In particular, we define $\varphi_{\mathcal{X}} = 0$ for $|\mathcal{X}| = 0$.

Based on the above congestion-taking assumption, we can define a CP-level congestion equilibrium as follows.

Definition *3:* A strategy profile $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$ is a CP-level congestion equilibrium of a game $(M, \mu, \mathcal{N}, s_I)$, if

$$(v_i - c)\rho_i\big(\varphi_{\mathcal{P}}\big) \begin{cases} \le v_i\ \rho_i\big(\varphi_{\mathcal{O}}\big), & \forall i \in \mathcal{O}, \\ > v_i\ \rho_i\big(\varphi_{\mathcal{O}}\big), & \forall i \in \mathcal{P}. \end{cases} \tag{6}$$

By comparing the two equilibrium solutions (5) and (6), we observe that the CP-level congestion equilibrium largely simplifies the Nash equilibrium. However, the congestion-taking assumption of the CPs might not be valid in practice if there exists CPs whose traffic do have "market power" and do substantially affect the congestion level of a service class unilaterally. If there is only one such market mover, we can again model the CPs' decision as a Stackelberg game, where the dominating CP will become the first-move, and the remaining CPs will move simultaneously afterwards. If there exists multiple market movers, it might be unavoidable to consider the inter-dependence among these players and adopt the complicated Nash equilibrium to capture the reality.

Theorem *4:* If $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$ is a CP-level congestion equilibrium of a game $(M, \mu, \mathcal{N}, s_I)$, then $\varphi_{\mathcal{P}} \leq \varphi_{\mathcal{O}}$.

Remark: Theorem 4 guarantees that the premium service class will be less congested than the ordinary service class under any CP-level congestion equilibrium. Interestingly, this is not always true under Nash equilibria.

Theorem *5 (Scaling of CP-level Congestion Equilibrium):* Under the scaling condition $\Phi(\xi\Lambda, \xi\mu) = \Phi(\Lambda, \mu), \forall \xi > 0$, if $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$ is an equilibrium of a game $(M, \mu, \mathcal{N}, s_I)$, it is also a same type of equilibrium (Nash or CP-level congestion equilibrium) of a game $(\xi M, \xi\mu, \mathcal{N}, s_I)$ for any $\xi > 0$.

Remark: Theorem 5 implies that if the congestion dynamics is independent of the scale of the system, the CP-level congestion equilibrium inherits the same property of the underlying user-level equilibrium (Theorem 3). Given this property, we characterize the CP-level equilibrium by only focusing on the per capita capacity μ/M of the system.

V. Numerical Algorithms and Applications

In this section, we first discuss the numerical methods and algorithms to practically solve different congestion equilibria introduced before. We will see that, compared to Nash equilibrium, there exists more efficient methods to solve the congestion equilibria. After that, we will discuss some extensions of our model and potential applications.

A. User-level Congestion Equilibrium

To solve a user-level congestion equilibrium φ for a system $(M, \mu, \mathcal{N})$, by Definition 1, we need to find the root of the equation $f(\phi) = \Phi(\Lambda(M, \phi), \mu) - \phi = 0$. By Assumption 1, we know that $\Lambda(M, \phi)$ is non-increasing in ϕ. By Assumption 2, we know that, given a fixed capacity, $\Phi(\Lambda, \mu)$ is non-decreasing in Λ. Thus, $\Phi(\Lambda(M, \phi), \mu)$ is non-increasing in ϕ and $f(\phi) = \Phi(\Lambda(\phi), \mu) - \phi$ is continuous and strictly decreasing in ϕ. As a result, we can evaluate the root of $f(\phi) = 0$ iteratively by using bisection search and converge to the equilibrium exponentially fast.

B. CP-level Congestion Equilibrium: A Special Case

The difficulty of finding a CP-level congestion equilibrium comes from the large number of CPs. Through exhaustive search, we need to evaluation 2^N possible scenarios. Here, we first introduce a condition under which we can efficiently obtain a CP-level congestion equilibrium and then explore a more general method to find CP-level congestion equilibria.

Assumption *6 (A Special Form of ρ_i):* For any CP $i \in \mathcal{N}$ and any congestion ϕ_1 and ϕ_2, $\rho_i(\phi_2) > 0\ \forall \phi < +\infty$ and

$$\frac{\rho_i(\phi_1)}{\rho_i(\phi_2)} = F_i(G(\phi_1, \phi_2)),$$

where $F_i(\cdot)$ is an increasing function, and $G(\phi_1, \phi_2)$ is a non-increasing in ϕ_1 and non-decreasing in ϕ_2.

The above assumption implies that there exist a common metric $G(\varphi_{\mathcal{O}}, \varphi_{\mathcal{P}})$ between the two levels of congestion in both service classes $\mathcal{O}$ and $\mathcal{P}$ that each CP i's achievable throughput ratio $\lambda_i(\varphi_{\mathcal{O}}) : \lambda_i(\varphi_{\mathcal{P}}) = \rho_i(\varphi_{\mathcal{O}}) : \rho_i(\varphi_{\mathcal{P}})$ can be evaluated as a function F_i of that common metric $G(\varphi_{\mathcal{O}}, \varphi_{\mathcal{P}})$. Two examples of ρ_i that satisfy the above assumption are

$$\rho_i(\phi) = \hat{\theta}_i e^{-\beta_i \phi}, \quad \text{and} \quad \rho_i(\phi) = \hat{\theta}_i (\beta_i)^{\phi} (\text{for } \beta_i \leq 1),$$

where β_i can serve as a differentiating parameter that reflects CP i's to sensitivity congestion. If the above two forms of ρ_i are adopted, the corresponding demand functions $d_i(\phi)$ under the proportional rate mechanism would be

$$d_i(\phi) = (\phi + 1) e^{-\beta_i \phi}, \quad \text{and} \quad d_i(\phi) = (\phi + 1)(\beta_i)^{\phi}.$$

Similarly, the corresponding demand functions $d_i(\varphi)$ under a max-min end-to-end congestion mechanism would be

$$d_i(\phi) = \max\{1, \hat{\theta}_i \phi\} e^{-\beta_i \phi}, \text{ and } d_i(\phi) = \max\{1, \hat{\theta}_i \phi\} (\beta_i)^{\phi}.$$

Under Assumption 6, we denote ζ_i as a relative priority of CP i under a given charge c, defined by

$$\zeta_i = \begin{cases} F_i^{-1}\Big(\dfrac{v_i - c}{v_i}\Big) & \text{if } v_i > c, \\ -\infty & \text{otherwise}. \end{cases} \tag{7}$$

The larger priority value a CP has, the larger chance it will end up in the premium class in an equilibrium.

Lemma *2:* Under Assumption 6, $s_{\mathcal{N}} = (\mathcal{O}, \mathcal{P})$ is a CP-level congestion equilibrium of a game $(M, \mu, \mathcal{N}, s_I)$, if and only if $\zeta_i \leq G(\varphi_{\mathcal{O}}, \varphi_{\mathcal{P}}) < \zeta_j$ for any $i \in \mathcal{O}$ and $j \in \mathcal{P}$.

Lemma 2 reveals the structure of a CP-level congestion equilibrium under which $\mathcal{O}$ always contains the CPs with smaller values of ζ_i. The following theorem further characterizes a condition under which a unique equilibrium exists.

Theorem *6:* Under Assumption 6, if the user-level congestion equilibrium satisfies $\varphi(M, \mu, \mathcal{N}') < \varphi(M, \mu, \mathcal{N})$ for all $\mathcal{N}' \subset \mathcal{N}$ and $\rho_i(\cdot)$ is strictly decreasing for all $i \in \mathcal{N}$, $(M, \mu, \mathcal{N}, s_I)$ has at most one CP-level congestion equilibrium.

Remark: Theorem 6 guarantees the uniqueness of a CP-level congestion equilibrium if it exists. If no equilibrium exists, there actually exists a CP that finds itself better off in $\mathcal{P}$ when it is in $\mathcal{O}$ and vice versa.

Based on Lemma 2 and Theorem 6, we propose the following bisection algorithm to calculate the unique CP-level congestion equilibrium for numerical evaluation purposes.

Bisection Algorithm for CP-Level Equilibrium

1. *Sort CPs based on ζ_i in an ascending order;*
2. l_1=0; $l_2 = N$; $l = \frac{1}{2}(l_1 + l_2)$;
3. **while** $l_2 > l_1$
4. $\quad \mathcal{H}_l = \{i \in \mathcal{N} : 1 \leq i \leq l\}$;
5. $\quad$ **if** $\zeta_l \leq G(\varphi_{\mathcal{H}_l}, \varphi_{\mathcal{N} \backslash \mathcal{H}_l}) < \zeta_{l+1}$ **then** $l_1 = l_2 = l$;
6. $\quad$ **if** $G(\varphi_{\mathcal{H}_l}, \varphi_{\mathcal{N} \backslash \mathcal{H}_l}) \geq \zeta_{l+1}$ **then** $l_1 = l$;
7. $\quad$ **if** $G(\varphi_{\mathcal{H}_l}, \varphi_{\mathcal{N} \backslash \mathcal{H}_l}) < \zeta_l$ **then** $l_2 = l$;
8. $\quad l = \frac{1}{2}(l_1 + l_2)$;
9. **return** $(\mathcal{H}_l, \mathcal{N} \backslash \mathcal{H}_l)$;

The bisection algorithm sorts the CPs based on their ζ_i values in an ascending order, and then, finds a partition of the sorted CPs under which the CPs with lower/higher indices will be in the service ordinary/premium class.

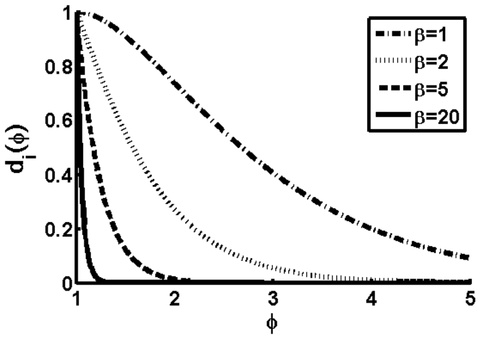


Fig. 2. Demand function $d_i(\phi)$.

C. CP-level Congestion Equilibrium: General Cases

For general forms of $\rho_i(\cdot)$s, we cannot sort the CPs on a one-dimensional space to search for a CP-level congestion equilibrium. However, we could still iteratively search in a two-dimensional "congestion space" as follows.

Iterative Congestion Taking Algorithm

1. *Initialize* $\varphi[0] = (\varphi_{\mathcal{O}}^{[0]}, \varphi_{\mathcal{P}}^{[0]})$
2. *Calculate induced equilibrium* $(\mathcal{O}_{[0]}, \mathcal{P}_{[0]})$ *given* $\varphi[0]$;
3. $t = 0$;
4. **do**
5. $\quad \varphi'[t] = (\varphi_{\mathcal{O}_{[t]}}, \varphi_{\mathcal{P}_{[t]}})$;
6. $\quad \varphi[t+1] = \varphi[t] + g[t](\varphi'[t] - \varphi[t])$;
7. $\quad t = t + 1$;
8. $\quad$ *Calculate* $(\mathcal{O}_{[t]}, \mathcal{P}_{[t]})$ *based on* $\varphi[t]$;
9. **until** $t > T$ *or* $(\mathcal{O}_{[t]}, \mathcal{P}_{[t]}) == (\mathcal{O}_{[t-1]}, \mathcal{P}_{[t-1]})$;
10. **return** $(\mathcal{O}_{[t]}, \mathcal{P}_{[t]})$;

The algorithm starts with an estimation $\varphi[0] = (\varphi_{\mathcal{O}}^{[0]}, \varphi_{\mathcal{P}}^{[0]})$ on the congestion levels in both service classes. At each iteration t, based on the congestion-taking assumption and the congestion estimation $\varphi[t]$, the algorithm calculate an induced CP-level congestion equilibrium $(\mathcal{O}_{[t]}, \mathcal{P}_{[t]})$ (line 2 and 8) and the corresponding real congestion level $\varphi'[t] = (\varphi_{\mathcal{O}_{[t]}}, \varphi_{\mathcal{P}_{[t]}})$ (line 5). The algorithm updates the congestion $\varphi[t+1]$ based on the previous estimation $\varphi[t]$ and the induced real congestion $\varphi'[t]$ (line 6). Notice that the algorithm is flexible that a step size $g[t]$ for each iteration (line 6) and a maximum number of iterations T (line 9) can be specified for the algorithm to tradeoff between convergence time and accuracy.

Remark: For the general cases, when there exist multiple CPs with similar characteristics, the system might have multiple CP-level of congestion equilibria. Similar to the special case, there might also be cases where no CP-level congestion equilibrium exists. However, our iterative algorithm provides a probable path of system dynamics, where CPs behave strategically based on observed levels of congestion.

D. User-level and CP-level Equilibria: Numerical Examples

We consider a proportional share mechanism of the system that induces system congestion according to Equation (2). We assume that the per-user achievable rate follows the function $\rho_i(\phi) = \hat{\theta}_i e^{-\beta_i \phi}$. Figure 2 illustrates the corresponding demand functions $d_i(\phi) = (\phi + 1)e^{-\beta_i \phi}$ with various values of β_i. We observe that the demand drops sharply with large β_is, which can be used to model content that have stringent throughput requirements, e.g. Netflix content. Content that is less sensitive to the throughput, e.g. a Google search query, can be modeled with a low β_i.

We first illustrate an example of user-level congestion equilibrium of two competing CPs in a system with a fixed user size $M = 1000$. We vary the available capacity μ from 0 to 4000. The two CPs have parameters $(\alpha_1, \hat{\theta}_1, \beta_1) = (0.05, 10, 5)$ and $(\alpha_2, \hat{\theta}_2, \beta_2) = (1, 1, 1)$. CP 1 represents Netflix-type of content that is more sensitive to congestion and has a higher unconstrained throughput rate; however, CP 2 represents Google-type of content that is more extensively used and less sensitive to congestion. Figure 3 illustrates the system congestion level φ, the throughput rates λ_i and the corresponding demands d_i of the two CPs from left to right. We observe that when μ increases from zero, the demand and throughput rate for congestion insensitive content increases sharply; while, the demand of congestion sensitive content does not start to catch up until the demand of congestion insensitive content reaches around 95%.

To visualize the CPs in CP-level congestion equilibria and get a better understanding of their behavior, we evaluate a scenario of 1000 CPs, whose α_i, $\hat{\theta}_i$ and v_i are uniformly distributed within $[0, 1]$. Each β_i is uniformly distributed within $[1, 10]$. Figure 4 illustrates eight equilibria under $c = 0.3$. The ISP choose κ to be either 0.2 or 0.9, the user size M is either 5×10^3 or 100×10^3, and the system capacity μ is either 1×10^6 or 20×10^6. CPs in the premium class are shown in red circles, which appear in the upper-right corner of the rectangle. We observe that 1) the set $\{i : v_i \leq 0.3, i \in \mathcal{N}\}$ is always in $\mathcal{O}$, 2) the same ratio of μ/M corresponds to the same equilibrium (Theorem 5), and 3) the number of CPs in the premium class, i.e. $|\mathcal{P}|$, increases against M and κ, but decreases against μ. This follows the intuition that more CPs move to the premium class when the system becomes more congested, due to a smaller ratio of μ/M, and/or a smaller capacity (large κ) for the ordinary class.

E. Extension and Applications

We develop the CP-level congestion equilibrium based on the underlying building block of user-level congestion equilibria. At the CP-level, the PMP-type of service differentiation of an ISP can naturally be extended from two service classes ($\mathcal{O}$ and $\mathcal{P}$) to many service classes. Our iterative algorithm for solving a CP-level congestion equilibrium can also be extended accordingly.

Although we only focused on a single eyeball ISP in the system, we can extend our model to a set $\mathcal{I}$ of ISPs, each $I \in \mathcal{I}$ of which uses its pricing and service differentiation strategy $s_I = (\kappa_I, c_I)$. Based on the congestion level and throughput rates in equilibrium, we can further derive the utilities, e.g., ISP revenue and CP utility, of different players in equilibrium. As an example, our work [15] focuses on throughput as a congestion metric and analyzes the consumer welfare under different ISP market structures. By comparing the resulting consumer welfare under congestion equilibria, we analyze the feasibility of network neutrality [26] regulations.

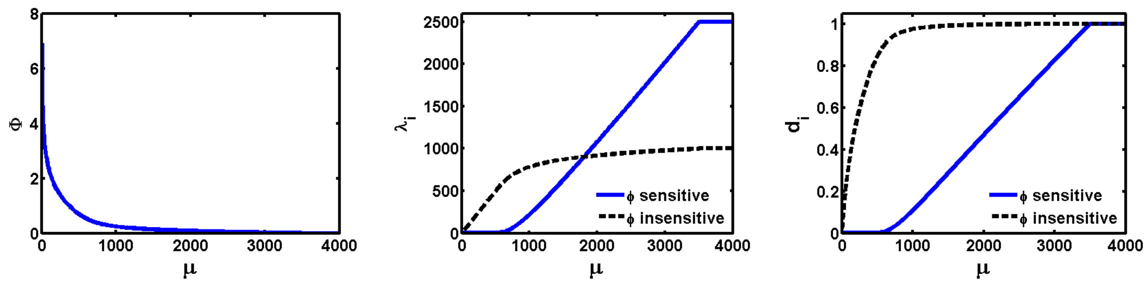


Fig. 3. System congestion, throughput rates and demands of two competing CPs.

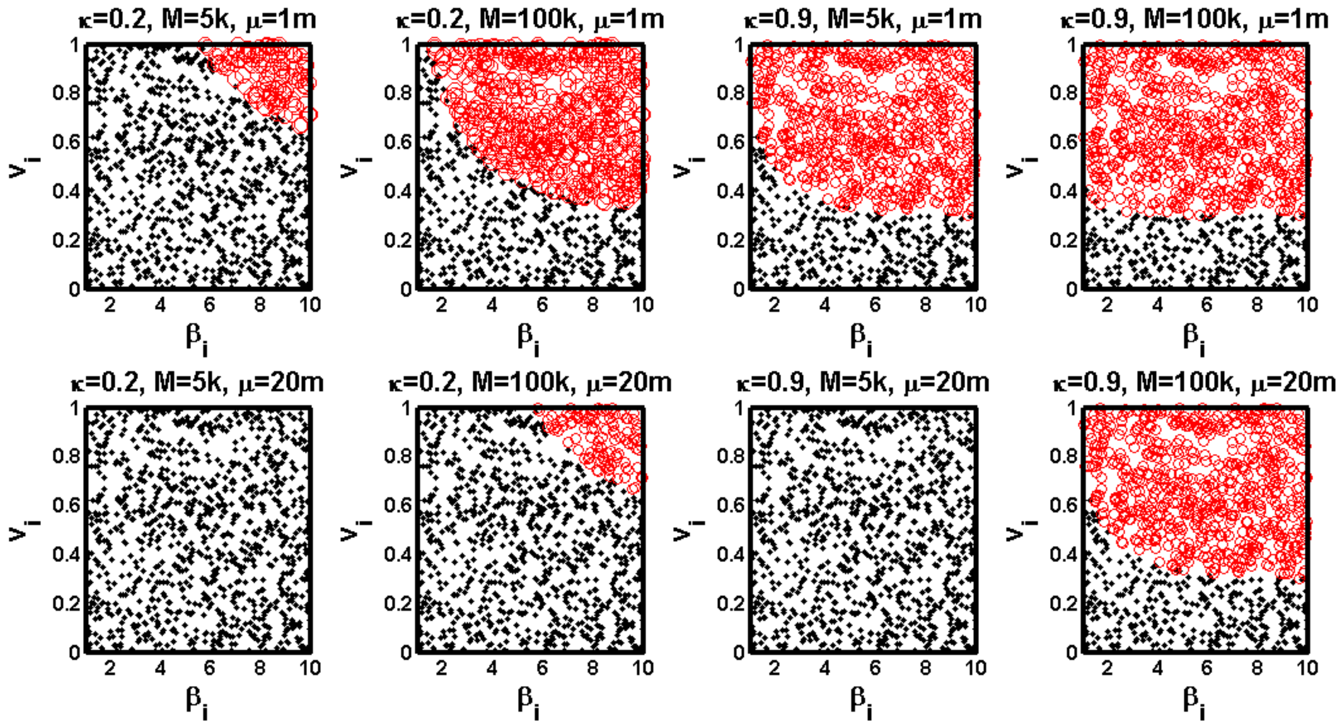


Fig. 4. CP-level congestion equilibria $(\mathcal{O}, \mathcal{P})$ under $c = 0.3$.

VI. Conclusions

In this paper, we have developed the concept of congestion equilibrium for large scale networks. The common experience of congestion by different players in a large scale network, coupled with the assumption that any individual player makes strategic decisions only based on the congestion level but not significantly affect the overall congestion, enables us to develop the concept of "congestion-taking". Under this congestion-taking assumption, we use the concept of congestion analogous to the price in traditional market economies and develop the congestion equilibrium of networks at different levels: the user-level and the CP-level. We have also applied our concept specifically to the study of modern multiparty Internet ecosystems comprising content producers, ISPs and end users, shedding new light on the topic and informing the network neutrality debate. Primarily, we believe that our ideas provide a new way to model and analyze modern network-economic systems that have a large number of decentralized, small players. Our work is a step in the way to help develop a better understanding of complex Internet ecosystems where engineering and economics are tightly coupled.

Acknowledgments

This study is supported by the research grant for the Human Sixth Sense Programme at the Advanced Digital Sciences Center from Singapore's Agency for Science, Technology and Research (A*STAR), Ministry of Education of Singapore AcRF grant R-252-000-448-133, and the National Science Foundation grants CNS-1017934 and CCF-1139915.

Appendix

Proof of Theorem 1: By Assumption 1, $\Lambda(M, \phi)$ is continuous and non-increasing in ϕ in the $\mathbb{R}_+^{|\mathcal{N}|}$ space. Further, by Assumption 2, $\Phi(\Lambda(M, \phi), \mu)$ is continuous and non-decreasing in ϕ. Let $f(\phi) = \Phi(\Lambda(M, \phi), \mu) - \phi$, which is a continuous and strictly decreasing function in ϕ. $f(\cdot)$ has a maximum value of $f(0) = \Phi(\Lambda(M, 0), \mu) > 0$

and $\lim_{\phi\to\infty} f(\phi) = -\infty$. Therefore, there exists a unique equilibrium φ that satisfies $f(\varphi) = 0$. ∎

Proof of Theorem 2: To simplify the notation, we define $\varphi_1 = \varphi(M,\mu_1,\mathcal{N}_1)$, $\varphi_2 = \varphi(M,\mu_2,\mathcal{N}_2)$, $\Lambda_1(M,\phi) = (\lambda_i(M,\phi) : i \in \mathcal{N}_1)$ and $\Lambda_2(M,\phi) = (\lambda_i(M,\phi) : i \in \mathcal{N}_2)$. Since $\mathcal{N}_1 \subseteq \mathcal{N}_2$, for any ϕ, $\Lambda_1(M,\phi)$ is a projection of $\Lambda_2(M,\phi)$ in the $\mathbb{R}_+^{|\mathcal{N}_1|}$ space. Suppose $\varphi_1 > \varphi_2$, by Assumption 1, we have $\Lambda_2(M,\varphi_1) \le \Lambda_2(M,\varphi_2)$. Further, by Assumption 2 and 3, we have $\Phi(\Lambda_1(M_1,\varphi_1),\mu_1) \le \Phi(\Lambda_1(M_2,\varphi_1),\mu_1) \le \Phi(\Lambda_2(M_2,\varphi_1),\mu_1) \le \Phi(\Lambda_2(M_2,\varphi_2),\mu_1) \le \Phi(\Lambda_2(M_2,\varphi_2),\mu_2)$. By Definition 1, $\varphi_1 = \Phi(\Lambda_1(M_1,\varphi_1),\mu_1) \le \Phi(\Lambda_2(M_2,\varphi_2),\mu_2) = \varphi_2$, which contradicts the assumption of $\varphi_1 > \varphi_2$. By Definition 1, Assumption 1 and the fact $\varphi_1 \le \varphi_2$, we conclude $\lambda_i(M,\mu_1,\mathcal{N}_1) = \lambda_i(M,\varphi_1) = \alpha_i M\rho_i(\varphi_1) \le \alpha_i M\rho_i(\varphi_2) = \lambda_i(M,\varphi_2) = \lambda_i(M,\mu_2,\mathcal{N}_2)$ for all $i \in \mathcal{N}_1$. ∎

Proof of Theorem 3: To simplify the notation, we define $\varphi_0 = \varphi(M,\mu,\mathcal{N})$ and $\varphi_\xi = \varphi(\xi M,\xi\mu,\mathcal{N})$. Suppose $\varphi_0 < \varphi_\xi$, by Assumption 1, we have $\Lambda(M,\varphi_0) \ge \Lambda(M,\varphi_\xi)$. By the CP's rate function of Equation (1), we have $\Lambda(\xi M,\varphi_\xi) = \xi\Lambda(M,\varphi_\xi)$ and therefore, $\Phi(\Lambda(\xi M,\varphi_\xi),\xi\mu) = \Phi(\xi\Lambda(M,\varphi_\xi),\xi\mu)$. Further, by Assumption 2 and 4, we have $\Phi(\Lambda(\xi M,\varphi_\xi),\xi\mu) = \Phi(\xi\Lambda(M,\varphi_\xi),\xi\mu) \le \Phi(\xi\Lambda(M,\varphi_0),\xi\mu) \le \Phi(\Lambda(M,\varphi_0),\mu)$. By Definition 1, the above implies that $\varphi_\xi = \Phi(\Lambda(\xi M,\varphi_\xi),\xi\mu) \le \Phi(\Lambda(M,\varphi_0),\mu) = \varphi_0$, which shows a contradiction. Therefore, we must have $\varphi_\xi \le \varphi_0$ and $\Lambda(M,\varphi_\xi) \ge \Lambda(M,\varphi_0)$ by Assumption 1. Finally, $\Lambda(\xi M,\xi\mu,\mathcal{N}) = \Lambda(\xi M,\varphi_\xi) = \xi\Lambda(M,\varphi_\xi) \ge \xi\Lambda(M,\varphi_0) = \xi\Lambda(M,\mu,\mathcal{N})$.

In particular, if $\Phi(\xi\Lambda,\xi\mu) = \Phi(\Lambda,\mu)$, we have $\varphi_\xi = \Phi(\Lambda(\xi M,\varphi_\xi),\xi\mu) = \Phi(\xi\Lambda(M,\varphi_\xi),\xi\mu) = \Phi(\Lambda(M,\varphi_\xi),\mu)$. Since $\varphi_\xi = \Phi(\Lambda(M,\varphi_\xi),\mu)$, φ_ξ is also a solution for the equilibrium congestion of system $(M,\mu,\mathcal{N})$. By Theorem 1, we have $\varphi_\xi = \varphi_0$ and $\Lambda(\xi M,\varphi_\xi) = \xi\Lambda(M,\varphi_\xi) = \xi\Lambda(M,\varphi_0)$. ∎

Proof of Lemma 1: The utility of CP i joining the premium and ordinary service classes are $(v_i - c)\ \lambda_i\big(M,\varphi_{\mathcal{P}\cup\{i\}}\big)$ and $v_i\lambda_i\big(M,\varphi_{\mathcal{O}\cup\{i\}}\big)$ respectively. By dividing the constant $\alpha_i M$ on both, we obtain the above condition. ∎

Proof of Theorem 4: Suppose $\varphi_\mathcal{P} > \varphi_\mathcal{O}$ for some CP-level congestion equilibrium $(\mathcal{O},\mathcal{P})$. By Assumption 1, we know $\rho_i(\varphi_\mathcal{P}) \le \rho_i(\varphi_\mathcal{O})$; therefore, $(v_i - c)\rho_i(\varphi_\mathcal{P}) \le (v_i - c)\rho_i(\varphi_\mathcal{O}) \le v_i\rho_i(\varphi_\mathcal{O})$. By Definition 3, we know $i \in \mathcal{O}$ for all CPs. This implies that $\mathcal{P} = \emptyset$ and $\varphi_\mathcal{P} = 0$ by Assumption 5. This contradicts our assumption of $\varphi_\mathcal{P} > \varphi_\mathcal{O} \ge 0$. ∎

Proof of Theorem 5: By Theorem 3, $(\mathcal{O},\mathcal{P})$ would induce the same level of congestion $\varphi_\mathcal{O}$ and $\varphi_\mathcal{P}$ in both service classes in a linearly scaled system $(\xi M,\xi\mu,\mathcal{N},s_I)$. Therefore, the right hand sides of both 5 and 6 do not change. Since the left hand sides of both 5 and 6 remain the same, all the equilibrium conditions are satisfied. ∎

Proof of Lemma 2: For any CP with $v_i \le c$, by Lemma 1, we know $i \in \mathcal{O}$. So, we check $-\infty = \zeta_i \le G(\varphi_\mathcal{O},\varphi_\mathcal{P})$. For any CP with $v_i > c$, by Assumption 6, inequality (6) becomes

$$\frac{v_i - c}{v_i}\begin{cases}\le \rho_i(\varphi_\mathcal{O})/\rho_i(\varphi_\mathcal{P}) = F_i(G(\varphi_\mathcal{O},\varphi_\mathcal{P})), & \forall i \in \mathcal{O},\\ > \rho_i(\varphi_\mathcal{O})/\rho_i(\varphi_\mathcal{P}) = F_i(G(\varphi_\mathcal{O},\varphi_\mathcal{P})), & \forall i \in \mathcal{P}.\end{cases}$$

Because $F_i(\cdot)$ is increasing, and therefore invertible, we have

$$\zeta_i = F^{-1}(\frac{v_i - c}{v_i})\begin{cases}\le G(\varphi_\mathcal{O},\varphi_\mathcal{P}), & \forall i \in \mathcal{O},\\ > G(\varphi_\mathcal{O},\varphi_\mathcal{P}), & \forall i \in \mathcal{P},\end{cases}$$

which is equivalent to $\zeta_i \le G(\varphi_\mathcal{O},\varphi_\mathcal{P}) < \zeta_j$ for all $i \in \mathcal{O}$ and $j \in \mathcal{P}$. ∎

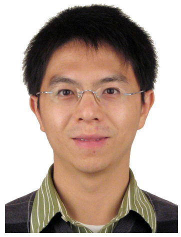
Proof of Theorem 6: We prove the uniqueness of CP-level congestion equilibrium by contradiction. Suppose there exist two equilibria $(\mathcal{O},\mathcal{P})$ and $(\mathcal{O}',\mathcal{P}')$ for some game $(M,\mu,\mathcal{N},s_I)$. By Lemma 2, we have $\zeta_i \le G(\varphi_\mathcal{O},\varphi_\mathcal{P}) < \zeta_j$ for any $i \in \mathcal{O}$ and $j \in \mathcal{P}$ and $\zeta_{i'} \le G(\varphi_{\mathcal{O}'},\varphi_{\mathcal{P}'}) < \zeta_{j'}$ for any $i' \in \mathcal{O}'$ and $j' \in \mathcal{P}'$. Therefore, if $|\mathcal{O}| = |\mathcal{O}'|$, we have $(\mathcal{O},\mathcal{P}) = (\mathcal{O}',\mathcal{P}')$. Without loss generality, we assume $|\mathcal{O}| < |\mathcal{O}'|$, which implies $\mathcal{O} \subset \mathcal{O}'$ and $\mathcal{P}' \subset \mathcal{P}$. Again by Lemma 2, we deduce that $G(\varphi_\mathcal{O},\varphi_\mathcal{P}) \ge G(\varphi_{\mathcal{O}'},\varphi_{\mathcal{P}'})$.

By the assumption $\varphi(M,\mu,\mathcal{N}') < \varphi(M,\mu,\mathcal{N})$ for all $\mathcal{N}' \subset \mathcal{N}$, we have $\varphi_\mathcal{O} < \varphi_{\mathcal{O}'}$ and $\varphi_\mathcal{P} > \varphi_{\mathcal{P}'}$. By the assumption that $\rho_i(\cdot)$ is strictly decreasing, $G(\phi_1,\phi_2)$ becomes strictly decreasing in ϕ_1 and strictly increasing in ϕ_2. As a result, we have $G(\varphi_\mathcal{O},\varphi_\mathcal{P}) < G(\varphi_{\mathcal{O}'},\varphi_{\mathcal{P}'})$, which shows a contradiction. ∎

REFERENCES

[1] Google corporate website. http://www.google.com/corporate/tech.html.
[2] Netflix performance on top isp networks. http://techblog.netflix.com/2011/01/netflix-performance-on-top-isp-networks.html.
[3] M. Armstrong. Competition in two-sided markets. *RAND Journal of Economics*, 37(3):668–691, 2006.
[4] R. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.
[5] D. Braess. ber ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12:258268, 1969.
[6] M. Campbell and J. Browning. Apple, Google asked to pay up as mobile operators face data flood. *Bloomberg News*, December 7 2010.
[7] X. Chen and X. Deng. Settling the complexity of 2-player Nash equilibrium. In *Proc. 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2006.
[8] D. M. Chiu and R. Jain. Analysis of the increase and decrease algorithms for congestion avoidance in computer networks. *Computer Networks and ISDN Systems*, 17(1).
[9] D. Clark. The design philosophy of the DARPA Internet protocols. *ACM SIGCOMM Computer Communication Review*, 18(4), August 1988.
[10] C. Courcoubetis and R. Weber. *Pricing Communication Networks: Economics, Technology and Modelling*. John Wiley & Sons Ltd., 2003.
[11] C. Daskalakis. On the complexity of approximating a Nash equilibrium. In *Proc. 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, San Francisco, California, USA, January 2011.
[12] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. In *J. Operational Research Society*, volume 49, 1998.
[13] F. P. Kelly. Charging and rate control for elastic traffic. *European Trans. Telecommunications*, 8:33–37, 1998.
[14] R. T. B. Ma, D. Chiu, J. C. Lui, V. Misra, and D. Rubenstein. On cooperative settlement between content, transit and eyeball Internet service providers. *IEEE/ACM Trans. Netw.*, 19(3), June 2011.
[15] R. T. B. Ma and V. Misra. The Public Option: a non-regulatory alternative to network neutrality. In *Proc. 7th ACM Conference on Emerging Network Experiment and Technology (CoNEXT 2011)*, Tokyo, Japan, December 2011.
[16] A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic theory*. Oxford University Press, 1995.

[17] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Netw.*, 8(5), October 2000.
[18] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14:124143, 1996.
[19] J. Nash. Equilibrium points in n-person games. *Proc. National Academy of Sciences*, 36(1):48–49, 1950.
[20] A. Odlyzko. Paris metro pricing for the Internet. *Proc. ACM Conference on Electronic Commerce*, 1999.
[21] M. J. Osborne and A. Rubinstein. *A course in game theory*. The MIT Press Cambridge, Massachusetts, 1994.
[22] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:6567, 1973.
[23] R. Selten. A reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.
[24] N. Shetty, G. Schwartz, and J. Walrand. Internet QoS and regulations. *IEEE/ACM Trans. Netw.*, 18(6), December 2010.
[25] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Institute of Civel Engineers*, 1(part 2):325–378, 1952.
[26] T. Wu. Network neutrality, broadband discrimination. *J. Telecommunications and High Technology Law*, 141, 2005.



Richard T. B. Ma received the B.Sc. (first-class honors) degree in Computer Science in July 2002 and the M.Phil. degree in Computer Science and Engineering in July 2004, both from the Chinese University of Hong Kong, and the Ph.D. degree in Electrical Engineering in May 2010 from Columbia University. During his Ph.D. study, he worked as a research intern at IBM T.J. Watson Research Center in New York and Telefonica Research in Barcelona. He is currently a Research Scientist in Advanced Digital Science Center, University of Illinois at Urbana-Champaign and an Assistant Professor in School of Computing at National University of Singapore. His current research interests include distributed systems, network economics, game theory, and stochastic processes.

Vishal Misra (S'98, M'99) received the B.Tech. degree from the Indian Institute of Technology, Bombay, India, in 1992, and the Ph.D. degree from the University of Massachusetts, Amherst, in 2000, all in Electrical Engineering.

He is an Associate Professor in Computer Science at Columbia University. He has received an NSF CAREER Award, a DoE CAREER Award and IBM Faculty Awards. His research emphasis is on mathematical modeling of computer systems, bridging the gap between practice and analysis. His recent work includes the areas of Internet Economics, Peer to Peer Networks and efficient Scheduling policies. He has served as the guest editor for the Journal of Performance Evaluation, TPC chair and General Chair of the ACM SIGMETRICS conference. He has participated as member of program committees for conferences such as IEEE Infocom, ACM SIGMETRICS, ACM SiGCOMM, IFIP Performance and IEEE ICNP and serves on the board of directors of ACM SIGMETRICS.