On the Price of Anarchy in Unbounded Delay Networks

Tao Wu
Nokia Research Center
Cambridge, Massachusetts, USA
tao.a.wu@nokia.com

David Starobinski
Boston University
Boston, Massachusetts, USA
staro@bu.edu

ABSTRACT

We investigate the worst case delay ratio between the Nash equilibrium and the social optimum in networks of N parallel links (routes) with unbounded delay functions. We compute this ratio, known as the "price of anarchy", for the case when the link delay functions correspond to $M/M/1$-FCFS or $M/G/1$-PS. For this problem, we find that the price of anarchy depends on the network topology in the sense that it is precisely equal to N. We then extend our results to $M/G/1$-FCFS and $G/G/1$-FCFS delay functions and compute the price of anarchy in a heavy load regime. Our results indicate that, even in very simple topological settings, the price of selfish behavior can potentially be very high.

1. INTRODUCTION

Distributed and selfish decision making is an essential aspect of many large scale decentralized systems such as the Internet [13]. As individual players selfishly compete for shared resources such as network bandwidth and CPU time, a fundamental question arises as how much worse the system performance gets (e.g., at Nash equilibrium), compared to that in a centrally controlled environment, where resources can be allocated optimally [6, 8]. This question, referred to as the price of anarchy, has attracted a great deal of research interest [16, 15, 5].

Past work on the price of anarchy has generally focused on bounded cost functions, which have finite cost as long as the load is finite. In the context of network routing, if the link delay (cost) function is affine, then the total cost at the Nash equilibrium of selfish routing, where no player has any incentive to individually deviate from its current strategy, is no worse than 4/3 times of the cost under an optimal allocation [16, 5, 14]. This result is independent of the network topology, e.g., it remains the same for two or more parallel links, as well as multi-commodity networks. More generally, for networks with bounded delay functions (that include constant delay functions), the price of anarchy is also independent of the topology [15].

GameNets '06, October 14, 2006, Pisa, Italy.

However, in most computer and communication networks such as in the Internet, link delays cannot simply be modeled using bounded or constant delay functions. Indeed, in these networks, the delay depends on the link congestion and consist of queueing, transmission, and propagation delay components. Thus, links delays have often been characterized using *unbounded* delay functions such as $M/M/1$ or $G/G/1$ queueing functions that are able to model the above delay components [11, 12, 4]. In this case, the relation between the price of anarchy and the network topology is much more difficult to determine. In [14], Roughgarden derived an upper bound on the price of anarchy for networks with $M/M/1$ and $M/G/1$ cost functions. This bound is useful in the case where the aggregate traffic load is smaller than the capacity of the slowest link in the network, otherwise it becomes infinite.

In this paper, we perform an exact analysis of the price of anarchy when the link cost function is unbounded. We investigate a simple network model that consists of two nodes connected by N parallel links. This model does not have to be interpreted literally, i.e., it can be used as an abstraction of practical network settings. For instance, the parallel links could represent physically disjoint paths. They also provide a good abstraction of overlay networks, such as content delivery networks and peer-to-peer networks, where the same service may be offered by multiple entities [17]. Other recent papers making use of the parallel link model in related applications include [2, 1, 7].

For this network model, we consider $M/M/1$ delay functions and derive the price of anarchy under arbitrary system load and link capacity configurations. We show that, in the worst-case, the delay ratio between the Nash equilibrium and the social optimum is precisely equal to N. We then extend our result to more general queueing models, such as $M/G/1$-FCFS and $G/G/1$-FCFS, and derive the price of anarchy at high traffic load. To the best of our knowledge, our result is among the first to show that the price of anarchy depends on the topology for networks with unbounded average delay. Compared with known results [14], this result is both *tight* (i.e., the lower and upper bounds match each other) and *general* (the result holds for any feasible load condition). One of our main observations is that as more links are added to the network, the worst-case inefficiency of selfish routing also increases.

The remainder of the paper is organized as follows. In Section 2, we describe our network model in detail. We establish fundamental bounds for price of anarchy for $M/M/1$ delay function in Section 3, and extend the results to $M/G/1$ and

$G/G/1$ in Section 4. We conclude the paper in Section 5.

2. THE MODEL

We consider a packet network consisting of a source vertex s, a destination vertex t, and N parallel links connecting them. The setting is a non-atomic game where infinitely many selfish agents independently generate traffic from s to t, each of whom has negligible contribution to the total traffic. The aggregate traffic arriving at s is a Poisson Process with rate λ packets/second. The service rate of link i is μ_i packets/second. Without loss of generality, we assume that the service rate vector $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_N]'$ is descending, i.e., $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_N$.

At source vertex s, each packet is independently assigned to link i with probability p_i. Note that p_i captures the overall probability of using link i, without reference to a specific individual agent. Thus, the arrival rate at link i is Poisson with rate $p_i\lambda$. We first consider an $M/M/1$ delay (or cost) function for link i:

$$\overline{T}_i(p_i) = \frac{1}{\mu_i - p_i\lambda}. \tag{1}$$

Let $\mathbf{p} = [p_1, p_2, \ldots, p_N]'$ denote the access probability vector. Then, for a given vector $\mathbf{p}$, the average delay of a packet in the network is

$$\overline{T}(\mathbf{p}) = \sum_{i=1}^{N} p_i \overline{T}_i(p_i) = \sum_{i=1}^{N} \frac{p_i}{\mu_i - p_i\lambda}. \tag{2}$$

The delay functions in Eq. (1) is important in telecommunications networks, because a variety of routing problems can be characterized using the $M/M/1$ model [4, 11]. However, the generality of this model extends beyond the network layer and can be found in the application layer as well. For example, the emergence of content replication networks such as content delivery networks and peer-to-peer networks can be effectively characterized using the $M/G/1$-Processing Sharing (PS) model, which shares the same delay function as Eq. (1) [17].

Note that any feasible vector $\mathbf{p}$ must satisfy several conditions. First, to be a legitimate probability vector, the coordinates of $\mathbf{p}$ must be non-negative and sum to one, i.e., $p_i \geq 0$ for all i and $\sum_{i=1}^{N} p_i = 1$. In addition, the coordinates of $\mathbf{p}$ must satisfy the *individual stability conditions*, i.e., $p_i < \mu_i/\lambda$, to guarantee that the arrival rate of requests is smaller than the service rate at each link i. We denote by $\mathcal{P}$ the set of vectors $\mathbf{p}$ that satisfy all the above constraints:

$$\mathcal{P} = \{\mathbf{p} \mid 0 \leq p_i < \mu_i/\lambda \quad \forall i, \quad \sum_{i=1}^{N} p_i = 1\}. \tag{3}$$

The set $\mathcal{P}$ is non-empty if and only if the *aggregate stability condition* $\lambda < \sum_{i=1}^{N} \mu_i$ is satisfied.

The social optimum (referred to as OPT hereafter) of the system is achieved when the vector $\mathbf{p}^*$ is found so that the average delay expression in Eq. (2) is minimized, subject to constraint $\mathcal{P}$ defined in Eq. (3):

$$\mathbf{p}^* = \arg\min_{\mathbf{p}\in\mathcal{P}} \overline{T}(\mathbf{p}).$$

Solving $\mathbf{p}^*$ for general delay functions has been the subject of studies in the literature in the context of load sharing in queueing networks [18, 10]. Theorem 1 below establishes a closed-form solution to the problem when the links have $M/M/1$ or $M/G/1$-PS delay functions, and also lays the theoretical foundation of this paper. Interested readers are referred to [17] for the derivation of the theorem.

THEOREM 1. *The optimal solution $\mathbf{p}^*$ to Eq. (2) under $M/M/1$ delay functions can be obtained as follows. Define*

$$\alpha_i = \frac{\mu_i}{\lambda} - \frac{\left(\sum_{j=1}^{i} \mu_j - \lambda\right)\sqrt{\mu_i}}{\lambda \sum_{j=1}^{i} \sqrt{\mu_j}} \qquad 0 \leq i \leq N. \tag{4}$$

Then,

1. *N^* is the maximum index i for which $\alpha_i > 0$.*
2. *$I(\mathbf{p}^*) = \{N^*+1, N^*+2, \ldots, N\}$,*
3. *$p_i^* = \frac{\mu_i}{\lambda} - \frac{\left[\sum_{j=1}^{N^*} \mu_j - \lambda\right]\sqrt{\mu_i}}{\lambda \sum_{j=1}^{N^*} \sqrt{\mu_j}} \quad \forall i \notin I(\mathbf{p}^*)$,*
4. *$p_i^* = 0 \quad \forall i \in I(\mathbf{p}^*)$.*
5. *$\overline{T}(\mathbf{p}^*)$ is computed via*

$$\begin{aligned} \overline{T}(\mathbf{p}^*) &= \sum_{i=1}^{N^*} \frac{p_i^*}{\mu_i - p_i^*\lambda} \\ &= \frac{\left(\sum_{i=1}^{N^*} \sqrt{\mu_i}\right)^2}{\lambda\left(\sum_{i=1}^{N^*} \mu_i - \lambda\right)} - \frac{N^*}{\lambda}. \end{aligned} \tag{5}$$

In Theorem 1, N^* can be interpreted as the number of active links (links with positive access probability) under OPT, and $I(\mathbf{p}^*)$ is the set of inactive links.

We now consider the delay at Nash equilibrium (referred to as NE hereafter), where an agent cannot improve the expected delay of its packets by changing its own access probability vector. The access probability vector at NE, $\hat{\mathbf{p}}$, must satisfy Wardrop's first and second principles, i.e., the average delays at all active links are the same and minimum [19, 15]. Formally, define

$$\mathcal{S} = \{\mathbf{p} \mid \overline{T}_i(p_i) = \overline{T}_j(p_j) \quad \forall i, j \notin I(\mathbf{p})\}, \tag{6}$$

then the NE access probability vector $\hat{\mathbf{p}}$ satisfies

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}\in(\mathcal{P}\cap\mathcal{S})} \left(\overline{T}(\mathbf{p})\right).$$

The following theorem, which is the analog of Theorem 1, provides closed-form solutions for $\hat{\mathbf{p}}$ and $\hat{N}$, the number of active links under NE [17]:

THEOREM 2. *The solution $\hat{\mathbf{p}}$ can be obtained as follows. Define*

$$\beta_i = \frac{\lambda + i\mu_i - \sum_{j=1}^{i} \mu_j}{i\lambda}, \qquad 0 \leq i \leq N. \tag{7}$$

Then,

1. *$\hat{N}$ is the maximum index i for which $\beta_i > 0$.*
2. *$I(\hat{\mathbf{p}}) = \{\hat{N}+1, \hat{N}+2, \ldots, N\}$,*
3. *$\hat{p}_i = \frac{\lambda + \hat{N}\mu_i - \sum_{j=1}^{\hat{N}} \mu_j}{\hat{N}\lambda} \quad \forall i \notin I(\hat{\mathbf{p}})$,*
4. *$\hat{p}_i = 0 \quad \forall i \in I(\hat{\mathbf{p}})$.*

5. $\overline{T}(\hat{\mathbf{p}})$ *is computed via*

$$\overline{T}(\hat{\mathbf{p}}) = \frac{\hat{N}}{\sum_{i=1}^{\hat{N}} \mu_i - \lambda}. \quad (8)$$

The purpose of this paper is to bound $\overline{T}(\hat{\mathbf{p}})/\overline{T}(\mathbf{p}^*)$ for any load λ, service vector $\boldsymbol{\mu}$, and number of links N.

3. BOUNDS WITH $M/M/1$ DELAY FUNCTIONS

We analytically determine the price of anarchy for $M/M/1$ link delay functions in this section. We will first develop a lower bound for the price of anarchy in Theorem 3, and then derive a matching upper bound in Theorem 4.

Before presenting our main results, we need the following lemma, which shows that the number of active links in NE is less than or equal to that in OPT. The proof can be found in [20].

LEMMA 1. *For any service rate vector $\boldsymbol{\mu}$, load λ and number of links N, the following relation holds:*

$$\hat{N} \leq N^* \leq N.$$

Our first result is that for a given number of links N, there exist arrival rate and service rate parameters such that the ratio $\overline{T}(\hat{\mathbf{p}})/\overline{T}(\mathbf{p}^*)$ approaches N. This situation occurs at high load, when $\hat{N} = N$.

THEOREM 3. *For a network with N parallel links,*

$$\sup_{\lambda, \boldsymbol{\mu}} \frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} \geq N.$$

PROOF. The theorem is proved by showing that for any given N, there exist parameters λ and μ_i, $i = 1, 2, \ldots, N$, such that $\overline{T}(\hat{\mathbf{p}})/\overline{T}(\mathbf{p}^*)$ becomes arbitrarily close to N, as done in Theorem 4 in [17]. $\square$

The main result of this paper is the following theorem. It offers a tight upper bound for $\overline{T}(\hat{\mathbf{p}})/\overline{T}(\mathbf{p}^*)$ when N^* is known.

THEOREM 4. *For a given system load λ and associated number of active links in OPT, N^*, the following inequality holds:*

$$\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} < N^*. \quad (9)$$

PROOF. From Lemma 1, we know $N^* \geq \hat{N}$ for any given λ. We prove the theorem in two steps. In the first step, we show that Eq. (9) holds when $N^* = \hat{N}$. In the second step, we construct two auxiliary systems to show that Eq. (9) holds when $N^* > \hat{N}$.

Step 1: The case of $N^* = \hat{N}$. We observe that in order for link $\hat{N}$ to be active, we must have $\beta_{\hat{N}} > 0$ or

$$\lambda + \hat{N}\mu_{\hat{N}} - \sum_{i=1}^{\hat{N}} \mu_i > 0. \quad (10)$$

We now define the *threshold service rate* μ_T as

$$\mu_T = \frac{\left(\sum_{i=1}^{\hat{N}} \mu_i - \lambda\right)}{\hat{N}}. \quad (11)$$

From Eq. (10), we can show that $\mu_{\hat{N}} > \mu_T$.

Because $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{\hat{N}}$, we have

$$\sqrt{\mu_i \mu_j} > \mu_T, \qquad \forall\ 1 \leq i, j \leq \hat{N}. \quad (12)$$

By summing both sides of Eq. (12) over all indices i between 1 and $\hat{N}$, we have

$$\hat{N}(\hat{N}-1)\mu_T < \sum_{i=1}^{\hat{N}} \sum_{\substack{j=1 \\ j \neq i}}^{\hat{N}} \sqrt{\mu_i \mu_j}. \quad (13)$$

We combine Eqs. (11) and (13) and obtain

$$\hat{N}^2 \mu_T + \lambda < \left(\sum_{i=1}^{\hat{N}} \sqrt{\mu_i} \right)^2. \quad (14)$$

Note that the derivation of Eq. (14) only assumes there are $\hat{N}$ active links in NE and does not require that $N^* = \hat{N}$. Hence it can be used in the next step as well.

Now we prove Eq. (9), when $N^* = \hat{N}$, which is equivalent to

$$\frac{\hat{N}\lambda}{\left(\sum_{i=1}^{\hat{N}} \sqrt{\mu_i}\right)^2 - \hat{N}\left(\sum_{i=1}^{\hat{N}} \mu_i - \lambda\right)} < \hat{N}. \quad (15)$$

Taking into account that $\mu_T = \frac{\left(\sum_{i=1}^{\hat{N}} \mu_i - \lambda\right)}{\hat{N}}$, we find that Eq. (15) is equivalent to Eq. (14).

Step 2: The case of $N^* > \hat{N}$. Our original system has N links with service rates $[\mu_1, \mu_2, \ldots, \mu_N]$. We first remove all inactive links in OPT from the original system. The resulting system, denoted as the $\mathbf{u}$ system, has a service rate vector

$$\mathbf{u} = [u_1, u_2, \ldots, u_{N^*}] = [\mu_1, \mu_2, \ldots, \mu_{\hat{N}}, \ldots, \mu_{N^*}].$$

Clearly, the $\mathbf{u}$ system's average delay under the OPT policy, $\overline{T}(\mathbf{p}^*_{\mathbf{u}})$, is the same as that of the original system, $\overline{T}(\mathbf{p}^*)$. Furthermore, according to Lemma 1, all removed links are inactive in NE too, so we have the same relation for NE: $\overline{T}(\hat{\mathbf{p}}_{\mathbf{u}}) = \overline{T}(\hat{\mathbf{p}})$. Consequently, the ratio $\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)}$ of the original system is the same as $\frac{\overline{T}(\hat{\mathbf{p}}_{\mathbf{u}})}{\overline{T}(\mathbf{p}^*_{\mathbf{u}})}$ of the $\mathbf{u}$ system.

The key of upper-bounding the delay ratio is constructing another auxiliary system, called the $\mathbf{v}$ system, which also has N^* links. In the $\mathbf{v}$ system, the fastest $\hat{N}$ links have the same rates as those in the $\mathbf{u}$ system, but the remaining $N^* - \hat{N}$ links are assigned the threshold service rate, μ_T. The service rate vector of the $\mathbf{v}$ system is thus

$$\mathbf{v} = [v_1, v_2, \ldots, v_{N^*}] = [\mu_1, \mu_2, \ldots, \mu_{\hat{N}}, \mu_T, \mu_T, \ldots, \mu_T]$$

Because $\mu_T < \mu_{\hat{N}}$, $\mathbf{v}$ is also descending. More importantly, the $\mathbf{u}$ and $\mathbf{v}$ systems have the same NE average delay, i.e., $\overline{T}(\hat{\mathbf{p}}_{\mathbf{u}}) = \overline{T}(\hat{\mathbf{p}}_{\mathbf{v}})$. This is because in the $\mathbf{v}$ system, at load λ, all links with the threshold service rate μ_T are inactive as shown in Theorem 2. Since the two systems have the same service rates for the active $\hat{N}$ links, they must have the same average delay when the NE policy is used.

Before we evaluate $\overline{T}(\mathbf{p}_{\mathbf{v}}^*)$, we need to show that all links in the $\mathbf{v}$ system are active under OPT. This is achieved by first observing from the derivation of N^* in Theorem 1, that in the $\mathbf{u}$ system,

$$\sqrt{u_{N^*}}\sum_{j=1}^{N^*}\sqrt{u_j} > \sum_{j=1}^{N^*} u_j - \lambda.$$

Because $u_{N^*-1} \geq u_{N^*}$, we have

$$\sqrt{u_{N^*-1}}\sum_{j=1}^{N^*-1}\sqrt{u_j} > \sum_{j=1}^{N^*-1} u_j - \lambda.$$

Repeating this procedure, we obtain

$$\sqrt{u_i}\sum_{j=1}^{i}\sqrt{u_j} > \sum_{j=1}^{i} u_j - \lambda, i = 1, 2, \ldots, N^*.$$

In particular,

$$\sqrt{u_{\hat{N}+1}}\sum_{j=1}^{\hat{N}+1}\sqrt{u_j} > \sum_{j=1}^{\hat{N}+1} u_j - \lambda$$

or

$$\sqrt{u_{\hat{N}+1}}\sum_{j=1}^{\hat{N}}\sqrt{u_j} > \sum_{j=1}^{\hat{N}} u_j - \lambda.$$

Because $v_{\hat{N}+1} = \mu_T \geq u_{\hat{N}+1}$, we have

$$\sqrt{v_{\hat{N}+1}}\sum_{j=1}^{\hat{N}}\sqrt{u_j} > \sum_{j=1}^{\hat{N}} u_j - \lambda$$

or

$$\sqrt{v_{\hat{N}+1}}\sum_{j=1}^{\hat{N}+1}\sqrt{v_j} > \sum_{j=1}^{\hat{N}+1} v_j - \lambda. \quad (16)$$

According to Eq. (4) in Theorem 1, Eq. (16) guarantees that link $\hat{N}+1$ is active under OPT in the $\mathbf{v}$ system. Because all links with indices larger than $\hat{N}$ in the $\mathbf{v}$ system have the same service rate, they must have the same access probability. Thus all links in the $\mathbf{v}$ system must be active under OPT.

On the other hand, the $\mathbf{v}$ system achieves the smallest OPT delay among all N^*-link systems whose fastest $\hat{N}$ links have the rates $\mu_1, \ldots, \mu_{\hat{N}}$ and who have $\hat{N}$ active links under NE under load λ. To see this, we observe that in all systems meeting the above requirements, we have $r_i \leq \mu_T$, for $\hat{N} < i \leq N^*$, where r_i is the service rate of the i-th link in any system under consideration; otherwise the system would have more than $\hat{N}$ active links under NE. Each link in the $\mathbf{v}$ system has the largest possible service rate among all systems meeting the above two requirements. As a result, it achieves the same or smaller average delay under OPT than any other system meeting the above two requirements. So we have $\overline{T}(\mathbf{p}_{\mathbf{v}}^*) \leq \overline{T}(\mathbf{p}_{\mathbf{u}}^*)$.

Combining the above analysis in both NE and OPT, we see that

$$\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} = \frac{\overline{T}(\hat{\mathbf{p}}_{\mathbf{u}})}{\overline{T}(\mathbf{p}_{\mathbf{u}}^*)} \leq \frac{\overline{T}(\hat{\mathbf{p}}_{\mathbf{v}})}{\overline{T}(\mathbf{p}_{\mathbf{v}}^*)}. \quad (17)$$

Now we upper bound $\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)}$ by evaluating $\overline{T}(\hat{\mathbf{p}}_{\mathbf{v}})$ and $\overline{T}(\mathbf{p}_{\mathbf{v}}^*)$.

By definition and the structure of $\mathbf{v}$, we know that

$$\hat{N}\mu_T = \sum_{i=1}^{\hat{N}}\mu_i - \lambda = \sum_{i=1}^{\hat{N}} v_i - \lambda$$

and

$$N^*\mu_T = \sum_{i=1}^{N^*} v_i - \lambda.$$

Hence, the $\mathbf{v}$ system's delay under NE is

$$\overline{T}(\hat{\mathbf{p}}_{\mathbf{v}}) = \overline{T}(\hat{\mathbf{p}}) = \frac{\hat{N}}{\sum_{i=1}^{\hat{N}} v_i - \lambda} = \frac{1}{\mu_T}. \quad (18)$$

For $\overline{T}(\mathbf{p}_{\mathbf{v}}^*)$, we have

$$\begin{aligned}\overline{T}(\mathbf{p}_{\mathbf{v}}^*) &= \frac{\left(\sum_{i=1}^{N^*}\sqrt{v_i}\right)^2}{\lambda\left(\sum_{i=1}^{N^*} v_i - \lambda\right)} - \frac{N^*}{\lambda} \\ &= \frac{\left[\sum_{i=1}^{\hat{N}}\sqrt{\mu_i} + (N^* - \hat{N})\sqrt{\mu_T}\right]^2 - (N^*)^2\mu_T}{\lambda N^*\mu_T}.\end{aligned}$$

Combining the expressions for $\overline{T}(\hat{\mathbf{p}}_{\mathbf{v}})$ and $\overline{T}(\mathbf{p}_{\mathbf{v}}^*)$ and incorporating Eq. (14), we have

$$\begin{aligned}\frac{\overline{T}(\hat{\mathbf{p}}_{\mathbf{v}})}{\overline{T}(\mathbf{p}_{\mathbf{v}}^*)} &< \frac{\lambda N^*}{\lambda + 2\left(N^* - \hat{N}\right)\sqrt{\mu_T}\left(\sum_{i=1}^{\hat{N}}\sqrt{\mu_i} - \hat{N}\sqrt{\mu_T}\right)} \\ &< N^*.\end{aligned}$$

From Eq. (17), we have

$$\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} < N^*,$$

which completes the proof of the second step. □

From Lemma 1 and Theorem 4, we obtain the following corollary.

Corollary 1. *The delay ratio between NE and OPT is upper bounded by the number of links in the system:*

$$\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} < N$$

Finally, combining Theorem 3 and Corollary 1, we have the following theorem.

Theorem 5. *The price of anarchy in networks composed of N parallel links with $M/M/1$ delay functions is N:*

$$\sup_{\lambda,\boldsymbol{\mu}}\frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} = N.$$

Theorem 5 is tight in the sense that the delay ratio of N is achievable with arbitrarily small error. Compared to previous results based on a bounded delay model [14, 15] (the price of anarchy is infinite unless the aggregate arrival rate is smaller than the lowest service rate), Theorem 5 is exact and applies to any feasible load condition. It also indicates that the worst-case inefficiency of selfish routing increases as more links are added to the network.

4. GENERALIZATIONS FOR $M/G/1$-FCFS AND $G/G/1$-FCFS AT HIGH LOAD

We first consider networks with $M/G/1$-FCFS delay functions, where the average delay consists of average service time $\frac{1}{\mu_i}$ and the average waiting time $\overline{W}_i(p_i)$, who is represented by the Pollaczek-Khinchin (PK) formula [11]:

$$\overline{W}_i(p_i) = \frac{K_i p_i \lambda}{\mu_i (\mu_i - p_i \lambda)}, \tag{19}$$

where $K_i \equiv \frac{1+C_S^2(i)}{2}$. $C_S(i)$ is defined based on the variance of the service time at link i, σ_i, and the i's service rate μ_i: $C_S(i) = \sigma_i \mu_i$. Thus, the link delay function $\overline{T}_i(p_i)$ is

$$\begin{aligned} \overline{T}_i(p_i) &= \frac{K_i p_i \lambda}{\mu_i (\mu_i - p_i \lambda)} + \frac{1}{\mu_i} \\ &= \frac{K_i}{\mu_i - p_i \lambda} + \frac{1 - K_i}{\mu_i}. \end{aligned} \tag{20}$$

It is generally infeasible to obtain a closed-form solution for $\overline{T}(\mathbf{p})$ when the link delay function is of the form in Eq. (20). While a variety of numerical methods can be applied to solve this problem [3, 18, 9], they do not offer much insight on the relative performance of NE compared to OPT. Here, we analyze the system behavior under high load, and are able to obtain results similar to those in Section 3.

Under high load, i.e., when $\frac{p_i \lambda}{\mu_i} \to 1$, for all $i = 1, 2, \ldots, N$, the overall delay is dominated by the waiting time (the first term on the right hand side of Eq. (20)):

$$\overline{T}_i(p_i) \cong \frac{K_i}{\mu_i - p_i \lambda} \tag{21}$$

Using the above delay function and following the same development procedure of Theorems 1, 2 and 3, we obtain the access probability for OPT and NE for $M/G/1$-FCFS link delay functions. We then derive the following lower bound on the price of anarchy, which is asymptotically tight at high load:

THEOREM 6. *For a network with N parallel links and delay function as in Eq. (21), we have*

$$\sup_{\lambda, \boldsymbol{\mu}} \frac{\overline{T}(\hat{\mathbf{p}})}{\overline{T}(\mathbf{p}^*)} \geq \frac{\sum_{i=1}^{N} K_i}{K_1}. \tag{22}$$

For $G/G/1$ delay functions, there does not exist simple and general expression for the average delays. However, in the heavy load regime, a simple, asymptotically exact expression for $\overline{W}_i$ can be provided. Specifically, define the coefficient of variation of the inter-arrival time at server i as $C_A(i)$, then [12]

$$\overline{W}_i(p_i) \cong \frac{C_A^2(i) + C_S^2(i)}{2 (\mu_i - p_i \lambda)}. \tag{23}$$

Using the notation $K_i = \frac{C_A^2(i)+C_S^2(i)}{2}$, we can directly apply Theorem 6 to lower bound $\overline{T}(\hat{\mathbf{p}})/\overline{T}(\mathbf{p}^*)$, for $G/G/1$-FCFS delay functions.

5. CONCLUSIONS

In this paper, we have examined the price of anarchy in a general unbounded delay network consisting of N parallel links. For $M/M/1$-FCFS and $M/G/1$-PS delay functions, we have proven an exact result directly linking the price of anarchy to the number of links N. This result differs from existing topology-independent results for bounded delay networks and indicates that the worst-case inefficiency of selfish routing increases with the size of the network. For the case of $G/G/1$-FCFS delay functions, we have been able to derive a lower bound (Eq. (22)) on the price of anarchy. This lower bound is tight at high load. We conjecture that the left hand side of Eq. (22) actually represents also an upper bound on the price anarchy for $G/G/1$-FCFS delay functions. The proof of this conjecture remains an open problem.

Acknowledgments

This research was supported in part by NSF CAREER grant CNS-0132802.

6. REFERENCES

[1] D. Acemoglu and A. Ozdaglar. Flow control, routing, and performance from service provider viewpoint. Technical Report WP1696, MIT-LIDS, May 2004.

[2] E. Altman, T. Basar, T. Jiménez, and N. Shimkin. Routing into two parallel links: Game-theoretic distributed algorithms. *Journal of Parallel and Distributed Computing*, 61(9):1367–1381, 2001.

[3] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, second edition, 1999.

[4] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, second edition, 1992.

[5] J. Correa, A. Schulz, and N. Stier-Moses. On the inefficiency of equilibria in congestion games. http://jcorrea.uai.cl/papers/CSS2005.pdf, 2005.

[6] D. Grosu and A. Chronopoulos. Noncooperative load balancing in distributed systems. *Journal of Parallel and Distributed Computing*, 65:1022–1034, 2005.

[7] A. Hayrapetyan, E. Tardos, and T. Wexler. A network pricing game for selfish traffic. In *PODC '05*, pages 284–291, Las Vegas, NV, 2005.

[8] H. Kameda, E. Altman, O. Pourtallier, J. Li, and Y. Hosokawa. Paradoxes in performance optimization of distributed systems. In *Proceedings of SSGRR 2000*, 2000.

[9] H. Kameda, J. Li, C. Kim, and Y. Zhang. *Optimal Load Balancing in Distributed Computer Systems*. Springer-Verlag, 1997.

[10] C. Kim and H. Kameda. An algorithm for optimal static load balancing in distributed computer systems. *IEEE Transactions on Computers*, 41(3):381–384, 1992.

[11] L. Kleinrock. *Queueing systems*, volume 1. Wiley, 1975.

[12] W. Marchal. Some simpler bounds on the mean queueing time. *Operations Research*, 26:1083–1088, 1978.

[13] C. Papadimitriou. Algorithms, games, and the Internet. In *STOC 2001*, 2001.

[14] T. Roughgarden. The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences*, 67(2):341–364, September 2003.

[15] T. Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.

[16] T. Roughgarden and E. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[17] D. Starobinski and T. Wu. Performance of server selection algorithms for content replication networks. In *NETWORKING 2005*, Waterloo, Ontario, May 2005.

[18] A. Tantawi and D. Towsley. Optimal static load balancing in distributed computer systems. *Journal of the ACM*, 32(2):445–465, 1985.

[19] J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings, Institution of Civil Engineers, PART II*, 1:325–378, 1952.

[20] T. Wu and D. Starobinski. Server selection for content replication networks: algorithms and analysis. Technical Report 2005-IR-0040, Center for Information and Systems Engineering, Boston University, Boston, MA, Aug. 2005.