

# The Value of Information in Selfish Routing

Simon Scherrer<sup>1</sup>, Adrian Perrig<sup>1</sup>, and Stefan Schmid<sup>2</sup>

<sup>1</sup> Department of Computer Science, ETH Zurich

<sup>2</sup> Faculty of Computer Science, University of Vienna

**Abstract.** Path selection by selfish agents has traditionally been studied by comparing social optima and equilibria in the Wardrop model, i.e., by investigating the Price of Anarchy in selfish routing. In this work, we refine and extend the traditional selfish-routing model in order to answer questions that arise in emerging path-aware Internet architectures. The model enables us to characterize the impact of different degrees of congestion information that users possess. Furthermore, it allows us to analytically quantify the impact of selfish routing, not only on users, but also on network operators. Based on our model, we show that the cost of selfish routing depends on the network topology, the perspective (users versus network operators), and the information that users have. Surprisingly, we show analytically and empirically that less information tends to lower the Price of Anarchy, almost to the optimum. Our results hence suggest that selfish routing has modest social cost even without the dissemination of path-load information.

**Keywords:** Price of Anarchy, Selfish Routing, Game Theory, Information

## 1 Introduction

If selfish agents are free to select communication paths in a network, their interactions can produce sub-optimal traffic allocations. A long line of research relating to *selfish routing* [27, 28, 30] has quantified many effects of distributed, uncoordinated path selection by selfish individuals in networks. While seminal work on such game-theoretic analyses dates back to Wardrop [37], especially the notion of *Price of Anarchy*, coined by Koutsoupias and Papadimitriou [17], has received much attention: the Price of Anarchy compares the worst possible outcome of individual decision making, i.e., the worst Nash equilibrium, to the global optimum, by taking the corresponding cost ratio. The Price of Anarchy in network path selection is typically measured in terms of *latency*.

In this paper, we revisit these concepts to investigate two key aspects which have been less explored in the literature so far and are highly relevant for newly emerging path-aware network architectures (cf. §1.1):

- ***Impact of information:*** A fundamental design question of network architectures concerns which information about the network state should be shared with end-hosts, beyond the latency information that can be observed by the end-hosts directly.

- **Impact on network operators:** While game-theoretic analyses usually revolve around the cost experienced by users, it is also important to understand the impact of selfish routing on the network operators' cost.

### 1.1 Practical Motivation

The traditional question studied in the selfish-routing literature, namely the efficiency of uncoordinated path selection by selfish agents, has recently received new relevance in the context of emerging Internet architectures relying on source-based path selection [2, 9, 25, 38]. In particular, the already deployed SCION architecture [3, 23] offers extensive path-selection control to users.

Today's Internet infrastructure is based on a forwarding mechanism that grants almost exclusive control to the network and almost no control to users (or *end-hosts*). In fact, all communication from a given end-host to another end-host takes place over the *single* AS-level path that BGP (Border Gateway Protocol) converged on. In the upcoming paradigm of *source-based path selection* [36], network operators supply end-hosts with a pre-selected set of paths to a destination, enabling end-hosts to select a forwarding path themselves.

Source-based path selection allows end-hosts to select paths with superior performance to the BGP-generated path [12, 16, 32] or to quickly switch to an alternative path upon link failures. However, a widely shared concern about source-based path selection regards the loss of control by network operators, which fear that the traffic distribution resulting from individual user decisions might impose considerable cost on both themselves and their customers. Another concern is that end-hosts require path-load information in order to perform path selection effectively, necessitating complex systems for the dissemination of network-state information. We refine and extend concepts from the selfish-routing literature to investigate the validity of these concerns.

### 1.2 Our Contributions

We present a game-theoretic model (§2) which allows us to quantify not only the Price of Anarchy experienced by end-hosts, but also to account for the network operators. Furthermore, we use our model to explore how end-host information about the network state affects the Price of Anarchy.

We find that different levels of information indeed lead to different Nash equilibria and thus also to different Prices of Anarchy. Intriguingly, we find that while more information can improve the efficiency of selfish routing in networks with few end-hosts (§3), more information tends to induce a *higher* Price of Anarchy in more general settings (§4). Indeed, near-optimal outcomes are typically achieved if end-hosts select paths based on simple latency measurements of different paths. These theoretical results suggest that source-based path selection cannot only achieve a good network performance in selfish contexts, but can be realized in a fairly light-weight manner, avoiding the need to distribute much information about the network state. This insight is validated with a case study on the Abilene topology (§5).

## 2 Model and First Insights

### 2.1 Model

As in previous work on selfish routing [10, 30], our model is inspired by the classic Wardrop model [37]. In this model, the network is abstracted as a graph  $G = (A, L)$ , where the edges  $\ell \in L$  between the nodes  $A_i \in A$  represent links. Every link  $\ell \in L$  is described by a link-cost function  $c_\ell(f_\ell)$ , where  $f_\ell$  is the amount of load on link  $\ell$ , i.e., a *link flow*. Typically, link-cost functions are seen as describing the latency behavior of a link. To reflect queuing dynamics, link cost functions are convex and non-decreasing. For every node pair  $(A_i, A_j)$ , there is a set of paths  $P(A_i, A_j)$  that contains all non-circular paths between  $A_i$  and  $A_j$ . Between any node pair  $(A_i, A_j)$ , there is a demand  $d$  shared by infinitely many agents, where each agent is controlling an infinitesimal share of traffic.

However, the traditional Wardrop model is not suitable to analyze traffic dynamics in an Internet context. We thus adapt the Wardrop model into a more realistic model as follows. First, we introduce the concept of *ASes* and *end-hosts*, which allows us to perform a more fine-grained analysis of traffic in an inter-domain network. An AS  $A_i \in A$  is represented by a node in the network graph  $G$ . The AS contains a set of end-hosts, which are the players in the path-selection game. Differently than in the Wardrop model, we allow for non-negligible, heterogeneous demand between end-host pairs in order to accommodate the variance of demand in the Internet. For example in origin-destination pair  $od = (e_s, e_t) \in OD$  (short:  $(s, t)$ ), an end-host  $e_s \in A_i$  can have a demand  $d_{s,t} \geq 0$  towards another end-host  $e_t \in A_j$ . We also deviate from the Wardrop model by considering a multi-path setting, where the demand  $d_{s,t}$  of one agent can be arbitrarily distributed over all paths  $p \in P(A_i, A_j)$ . The amount of flow from end-host  $e_s$  to end-host  $e_t$  on path  $p \in P(A_i, A_j)$  is denoted as a *path flow*  $F_{(s,t),p}$ , which must be non-negative, with  $\sum_{p \in P(A_i, A_j)} F_{(s,t),p} = d_{s,t}$ . The set  $\Pi(e_s, e_t) \subseteq \Pi$  contains all end-host paths of the form  $\pi = [(s, t), p]$ , where  $e_s, e_t$  are end-hosts connected by the AS-level path  $p$ . All path flows  $F_{(s,t),p}$  for an origin-destination pair  $(e_s, e_t)$  are collected in a path-flow vector  $\mathbf{F}_{s,t} \in \mathbb{R}^{|\Pi(e_s, e_t)|}$ . All such path-flow vectors  $\mathbf{F}_{s,t}$  are collected in the global *path-flow pattern*  $\mathbf{F} \in \mathbb{R}^{|\Pi|}$ . A link flow  $f_\ell$  for link  $\ell \in L$  is the sum of the path flows in  $\mathbf{F}$  that refer to end-host paths  $\pi$  containing link  $\ell$ , i.e.,  $f_\ell = \sum_{\pi \in \Pi: \ell \in \pi} F_\pi$ .

The cost  $C_\pi$  of an end-host path  $\pi$  given a certain path-flow pattern  $\mathbf{F}$  is the sum of the cost of all links in the path:  $C_\pi(\mathbf{F}) = \sum_{\ell \in \pi} c_\ell(f_\ell)$ . The cost to end-hosts  $C^*(\mathbf{F})$  from a path-flow pattern  $\mathbf{F}$  is the latency experienced by all end-hosts on all the paths to all of their destinations, weighted by the amount of traffic that goes over a given path. This term can be simplified as follows:

$$C^*(\mathbf{F}) = \sum_{(s,t) \in OD} \sum_{\pi \in \Pi(s,t)} F_\pi \cdot C_\pi(\mathbf{F}) = \sum_{\pi \in \Pi} F_\pi \cdot \sum_{\ell \in \pi} c_\ell(f_\ell) = \sum_{\ell \in L} f_\ell \cdot c_\ell(f_\ell)$$

Existing work on selfish routing [28, 30] usually defines total cost in the above sense. However, when analyzing source-based path selection architectures, the

network-operator perspective on social cost is essential. Therefore, we also introduce a social cost function relating to the perspective of network operators.

The basic idea of the network-operator cost function  $C^\#$  is to treat links as investment assets. Thus, the business performance of a link  $\ell$  is given by a function  $p_\ell^\#(f_\ell) = b_\ell^\#(f_\ell) - c_\ell^\#(f_\ell)$ , where  $b_\ell^\#$  and  $c_\ell^\#$  are the benefits and costs of a link, respectively. As we investigate effects on the aggregate of network operators, we model the network-operator cost function as follows:

$$C^\#(\mathbf{F}) = \sum_{\ell \in L} -p_\ell^\#(f_\ell) = \sum_{\ell \in L} c_\ell^\#(f_\ell) - b_\ell^\#(f_\ell) = \sum_{\ell \in L} c_\ell(f_\ell)$$

We justify this formulation as follows. Concerning link costs  $c_\ell^\#$ , a central insight is that network-operator costs mostly stem from heavily used links. In volume-based interconnection agreements, excessive usage of a link induces high charges, whereas in peering agreements, excessive usage violates the agreement and triggers expensive renegotiation. Moreover, heavy usage necessitates expensive capacity upgrades. As the latency function  $c_\ell(f_\ell)$  indicates the congestion level on link  $\ell$ , we approximate  $c_\ell^\# \approx c_\ell$ . The link benefit  $b_\ell^\#$  captures the link revenue, both revenue from customer ASes and customer end-hosts. In the aggregate, the monetary transfers between ASes (charges paid and received) sum up to zero. Given a fixed market size, the revenue from end-hosts sums up to a constant in the aggregate. Hence, the global benefit  $\sum_{\ell \in L} b_\ell^\#$  is constant and can be dropped, as the absolute level of the network-operator cost is irrelevant for our purposes. This convex formulation of  $C^\#$  allows theoretical analysis.

## 2.2 Social Optima

According to Wardrop [6,37], a socially optimal traffic distribution is reached iff the total cost cannot be reduced by moving traffic from one path to another. In the optimum, the cost increase on an additionally loaded path at least outweighs the cost reduction from a relieved path. Because the cost functions are convex and non-decreasing, it suffices that this condition holds for an infinitesimal traffic share. Adding an infinitesimal amount to the argument of a cost function imposes a *marginal cost*, given by the derivative of the cost function. A socially optimal traffic distribution is thus reached iff the marginal cost of every alternative path is not smaller than the marginal cost of the currently used paths [6]:

**Social optimum.** A path-flow pattern  $\mathbf{F}$  represents a social optimum w.r.t. cost function  $C$  if and only if for every origin-destination pair  $od \in OD$ , the paths  $\pi_1, \dots, \pi_i, \pi_{i+1}, \dots, \pi_{|\Pi(od)|} \in \Pi(od)$  stand in the following relationship:

$$\frac{\partial}{\partial F_{\pi_1}} C(\mathbf{F}) = \dots = \frac{\partial}{\partial F_{\pi_i}} C(\mathbf{F}) \leq \frac{\partial}{\partial F_{\pi_{i+1}}} C(\mathbf{F}) \leq \dots \leq \frac{\partial}{\partial F_{\pi_{|\Pi(od)|}}} C(\mathbf{F})$$

$$F_\pi > 0 \quad \text{for } \pi = \pi_1, \dots, \pi_i, \quad F_\pi = 0 \quad \text{for } \pi = \pi_{i+1}, \dots, \pi_{|\Pi(od)|}.$$

In this work, we refine the conventional notion of the social optimum by distinguishing two different perspectives on social cost: The *end-host optimum*  $\mathbf{F}^*$  satisfies the above conditions with respect to the function  $C^*$ , whereas the *network-operator optimum*  $\mathbf{F}^\#$  satisfies the above conditions with respect to function  $C^\#$ .

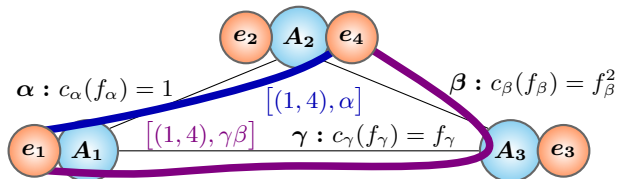


Fig. 1: Example network illustrating the source-based path selection model.

Interestingly, the end-host optimum  $\mathbf{F}^*$  and the network-operator optimum  $\mathbf{F}^\#$  can differ substantially. Assume that end-host  $e_1$  in Figure 1 has a demand of  $d_{1,4} = 1$  towards end-host  $e_4$  and that there is no other traffic in the network. The network-operator cost function  $C^\#(\mathbf{F})$  is  $1 + F_{\gamma\beta}^2 + F_{\gamma\beta}$  and is minimized by  $\mathbf{F}^\# = (1, 0)^\top$ , i.e., by sending all traffic over link  $\alpha$ . In contrast, the end-host cost function is  $F_\alpha + F_{\gamma\beta}^3 + F_{\gamma\beta}^2$  and is minimized by  $\mathbf{F}^* = (2/3, 1/3)^\top$ , i.e., by sending two thirds of traffic over link  $\alpha$  and the remaining third over path  $\gamma\beta$ .

### 2.3 Degrees of Information

In this paper, we consider the following two assumptions on the network information possessed by end-hosts:

- **Latency-only information (LI):** End-hosts know the latency of every path to a destination.
- **Perfect information (PI):** End-hosts know not only the latency of different paths, but also how the latency of the network links depends on the current load, i.e., the *latency functions*. Moreover, the end-hosts know the current link utilization, i.e., the background traffic.

The LI assumption hence reflects a scenario where end-hosts have to rely solely on latency measurements of paths, i.e., through RTT measurements from their own device. The LI assumption is the standard model traditionally considered in the selfish routing literature [11, 17, 30].

In this work, we extend the standard model by introducing the concept of *perfect information* (PI). The PI assumption reflects a scenario where end-hosts can always take the *best* traffic-allocation decision in selfish terms. More specifically, the PI assumption allows end-hosts to compute the *marginal cost* of a path. In path-aware networking, supplying end-hosts with perfect information is possible, as such information is known by network operators and can be disseminated along with path information.

Figure 2 illustrates the difference between the LI assumption and the PI assumption. Assume that end-host  $e$ , residing in AS  $A$ , has a demand of  $d = 1$

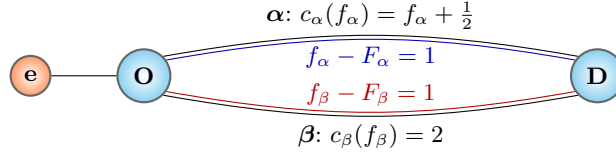


Fig. 2: Example illustrating the different degrees of end-host information.

to a destination in AS  $D$ . End-host  $e$  can split its traffic between two paths  $\alpha$  and  $\beta$ , both consisting of a single link with the cost functions  $c_\alpha$  (linear) and  $c_\beta$  (constant). The background traffic (traffic not from end-host  $e$ ) is 1 on both paths. Assuming the traffic allocation of end-host  $e$  is  $(F_\alpha, F_\beta) = (0.5, 0.5)$ , the path-latency values are given by  $c_\alpha(0.5+1) = 2$  and  $c_\beta(0.5+1) = 2$ . Given the LI assumption, end-host  $e$  performs no traffic reallocation, as there is no lower-cost alternative path which traffic could be shifted to. Moreover, there is no method for predicting the path costs for a different traffic allocation. However, such a prediction is possible with perfect information (PI): under the PI assumption, end-host  $e$  knows the cost functions and the background traffic such that it can optimize the objective  $F_\alpha \cdot (F_\alpha + 1 + \frac{1}{2}) + (1 - F_\alpha) \cdot 2$ . As a result, end-host  $e$  discovers the optimal traffic assignment  $(0.25, 0.75)$ . Intriguingly, the more detailed perfect information (PI) enables end-host  $e$  to detect an optimization that it cannot directly observe when confronted with latency values only (LI).

## 2.4 Nash Equilibria

In general, uncoordinated actions of selfish end-hosts do not result in socially optimal traffic allocations. Instead, the only stable states that arise in selfish path selection are *Nash equilibria*, i.e., situations in which no end-host perceives an opportunity to reduce its selfish cost by unilaterally reallocating traffic. However, as shown in §2.3, the degree of available information (LI or PI) strongly influences the optimization opportunities that an end-host perceives. Therefore, different information assumptions induce different types of Nash equilibria:

**LI equilibrium.** An end-host restricted to latency measurements will shift traffic from high-cost paths to low-cost paths whenever there is a cost discrepancy between paths, and will stop reallocating traffic whenever there is no lower-cost path anymore which the traffic could be shifted to. In the latter situation, an end-host under the LI assumption cannot perceive any way of reducing its selfish cost. A Nash equilibrium under the LI assumption (short: LI equilibrium) can thus be defined as follows:

**LI equilibrium.** A path-flow pattern  $\mathbf{F}$  represents an LI equilibrium  $\mathbf{F}^0$  if and only if for every origin-destination pair  $od \in OD$ , the paths  $\pi_1, \dots, \pi_i, \pi_{i+1}, \dots, \pi_{|\Pi(od)|} \in \Pi(od)$  have the following relationship:

$$C_{\pi_1}(\mathbf{F}) = \dots = C_{\pi_i}(\mathbf{F}) \leq C_{\pi_{i+1}}(\mathbf{F}) \leq \dots \leq C_{\pi_{|\Pi(od)|}}(\mathbf{F})$$

$$F_\pi > 0 \quad \text{for } \pi = \pi_1, \dots, \pi_i \quad \quad \quad F_\pi = 0 \quad \text{for } \pi = \pi_{i+1}, \dots, \pi_{|\Pi(od)|}$$

Traditionally, selfish-routing literature [11, 27, 30] considers a Nash equilibrium in the sense of the LI equilibrium, namely an equilibrium defined by the cost equality of all used paths to a destination. Under this classical definition, selfish routing is an instance of a *potential game* [31].

**PI Equilibrium.** We contrast the classical equilibrium (LI equilibrium) with a different equilibrium definition that builds on our new concept of perfect information (PI). As explained in §2.3, the PI assumption states that end-hosts do not only possess cost information of available paths to a destination, but are informed about the cost *functions* of all links in the available paths, as well as the background traffic on these links, i.e., the arguments to the cost functions. An end-host can thus calculate the selfish cost of a specific traffic reallocation and find the path-flow pattern that minimizes the end-host’s selfish cost.

The selfish cost  $C_{(e)}^*(\mathbf{F})$  of end-host  $e$  is given by the cost of all paths to all desired destinations, weighted by the amount of flow relevant to end-host  $e$ :

$$C_{(e)}^*(\mathbf{F}) = \sum_{\ell \in L} f_{\ell,(e)} \cdot c_{\ell}(f_{\ell})$$

where  $f_{\ell,(e)}$  is the flow volume on link  $\ell$  for which  $e$  is origin or destination.

Similar to the end-host social cost function  $C^*$  of which it is a partial term,  $C_{(e)}^*$  has a minimum that is characterized by a marginal-cost equality. An equilibrium under the PI assumption is thus given if and only if all end-hosts are at the minimum of their respective selfish cost functions, given the traffic by all other end-hosts:

**PI equilibrium.** A path-flow pattern  $\mathbf{F}$  represents a PI equilibrium  $\mathbf{F}^+$  if and only if for every origin-destination pair  $od = (e, \_) \in OD$ , the paths  $\pi_1, \dots, \pi_i, \pi_{i+1}, \dots, \pi_P \in \Pi(od)$  stand in the following relationship:

$$\frac{\partial}{\partial F_{\pi_1}} C_{(e)}^*(\mathbf{F}) = \dots = \frac{\partial}{\partial F_{\pi_i}} C_{(e)}^*(\mathbf{F}) \leq \frac{\partial}{\partial F_{\pi_{i+1}}} C_{(e)}^*(\mathbf{F}) \leq \dots \leq \frac{\partial}{\partial F_{\pi_{|\Pi(od)|}}} C_{(e)}^*(\mathbf{F})$$

$$F_{\pi} > 0 \quad \text{for } \pi = \pi_1, \dots, \pi_i \qquad F_{\pi} = 0 \quad \text{for } \pi = \pi_{i+1}, \dots, \pi_{|\Pi(od)|}$$

## 2.5 Price of Anarchy

A natural way of analyzing the efficiency of selfish routing is to compare the social optima and the equilibria in a network. Typically, such a comparison involves computing the *Price of Anarchy (PoA)*, i.e., the ratio of the equilibrium cost and the optimal cost. By definition of the optimal cost, this ratio is always larger or equal to 1.

In our model, the classical Price of Anarchy from the existing literature reflects a comparison of the end-host cost  $C^*$  of the LI equilibrium  $\mathbf{F}^0$  and the end-host cost  $C^*$  of the end-host optimum  $\mathbf{F}^*$ . With the additional versions of social optima and equilibria established in the preceding sections, a total of four different variants of the Price of Anarchy are possible, one for each combination of equilibrium (LI or PI) and perspective (end-hosts or network operators):

	LI equilibrium	PI equilibrium
End-host perspective	$PoA^{*0} = \frac{C^*(\mathbf{F}^0)}{C^*(\mathbf{F}^*)}$	$PoA^{*+} = \frac{C^*(\mathbf{F}^+)}{C^*(\mathbf{F}^*)}$
Network-operator perspective	$PoA^{\#0} = \frac{C^\#(\mathbf{F}^0)}{C^\#(\mathbf{F}^\#)}$	$PoA^{\#+} = \frac{C^\#(\mathbf{F}^+)}{C^\#(\mathbf{F}^\#)}$

Table 1: Different versions of the Price of Anarchy.

## 2.6 Value of Information

To compare different equilibria for different information assumptions, we introduce the *Value of Information (VoI)*. For a given perspective, the Value of Information is the difference between the Prices of Anarchy under the LI and PI assumptions, denominated by the Price of Anarchy under the LI assumption:

$$VoI^* = \frac{PoA^{*0} - PoA^{*+}}{PoA^{*0}} \quad VoI^\# = \frac{PoA^{\#0} - PoA^{\#+}}{PoA^{\#0}}$$

A positive Value of Information reflects a situation where the equilibrium under the PI assumption is closer to the social optimum than the equilibrium under the LI assumption. We identify and analyze scenarios with a positive impact of information in §3. A negative Value of Information reflects the counter-intuitive scenario where additional information makes the equilibrium more costly (cf. §4).

## 3 The Benefits of Information

In this section, we will show that information is beneficial in the artificial network settings traditionally considered in the literature [27]. More precisely, we show that in this setting, the PI equilibrium induces a lower Price of Anarchy than the LI equilibrium such that the Value of Information is positive. This is intuitive: if end-hosts possess more information, source-based path selection is more efficient.

In the network of Figure 3,  $K$  end-hosts  $e_1, \dots, e_K$  reside in AS  $O$ . Each end-host has a demand of  $d/K$  towards a destination in AS  $D$ . ASes  $O$  and  $D$  are connected by  $m$  links  $\alpha_1, \dots, \alpha_m$  with a constant cost function  $c_{\alpha_i}(f_{\alpha_i}) = d^p$  and one link  $\beta$  with a load-dependent cost function  $c_\beta(f_\beta) = f_\beta^p$ , where  $p \geq 1$ .

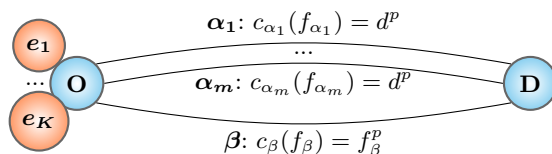


Fig. 3: Example network with beneficial impact of end-host information.

Such networks of parallel links are of special importance in the theoretical selfish-routing literature. In particular, Roughgarden [27] proved that the network in Figure 3 reveals the worst-case Price of Anarchy for any network with link cost functions limited to polynomials of degree  $p$ . The intuition behind this result is that the Price of Anarchy relates to a difference of steepness between



cost functions of competing links: the link  $\beta$  allows to reduce the cost of traffic from AS  $O$  to AS  $D$  if used modestly, but loses its advantage over the links  $\alpha_i$  if fully used. However, in selfish routing, end-hosts will use link  $\beta$  until the link is fully used, as it is always a lower-cost alternative path if not fully used. Therefore, the end-hosts overuse link  $\beta$  compared to the optimum. Intuitively, the parallel-links network represents a network where end-hosts have a choice between paths with different latency behavior.

Roughgarden's result refers to the classical Price of Anarchy, i.e., the Price of Anarchy  $PoA^{*0}$  to end-hosts under the LI assumption. In this section, we will show how this result is affected by additionally introducing the network-operator perspective and the PI assumption. In particular, we will prove the following theorem:

**Theorem 1.** *In a network of parallel links, a higher degree of information (PI assumption) is always more socially beneficial compared to a lower degree of information (LI assumption), both from the perspective of end-hosts and network operators:*

$$PoA^{*+} \leq PoA^{*0} \quad PoA^{\#+} \leq PoA^{\#0}$$

### 3.1 Social Optima

The end-host optimum in the network of parallel links can be shown to have social cost  $C^*(\mathbf{F}^*) = d^{p+1}[1 - p/(p+1)^{(p+1)/p}]$ . As the derivation is relatively similar to Roughgarden [27], it has been moved to Appendix A.1.

The network-operator optimum  $\mathbf{F}^\#$  is simple to derive: Since the cost of the links  $\alpha_i$  is independent of the flow on these links in contrast to the cost of link  $\beta$ , any flow on link  $\beta$  increases the cost  $C^\#$  to network operators. The minimal cost to network operators is thus simply  $C^\#(\mathbf{F}^\#) = m \cdot d^p$ .

### 3.2 LI Equilibrium

Under the LI assumption, a network is in equilibrium if for every end-host pair, all used paths have the same cost and all unused paths do not have a lower cost. Applied to the simple network in Figure 3, this condition is satisfied if and only if  $f_\beta^0 = d$  and  $f_{\alpha_i}^0 = 0 \forall f_{\alpha_i}$ , implying  $c_\beta(f_\beta^0) = d^p = c_{\alpha_i}(f_{\alpha_i}^0)$ . The path-flow pattern  $\mathbf{F}^0$  with  $F_{(k,D),\beta} = d/K$  and  $F_{(k,D),\alpha_i} = 0$  represents the LI equilibrium. The cost  $C^*$  of the LI equilibrium  $\mathbf{F}^0$  to end-hosts is simply  $C^*(\mathbf{F}^0) = d^{p+1}$ . The Price of Anarchy to end-hosts under the LI assumption is thus  $PoA^{*0} = C^*(\mathbf{F}^0)/C^*(\mathbf{F}^*) = [1 - p/(p+1)^{(p+1)/p}]^{-1}$ .

The cost  $C^\#$  of the LI equilibrium  $\mathbf{F}^0$  to network operators is given by  $C^\#(\mathbf{F}^0) = d^p + \sum_{\alpha_i} d^p = (m+1) \cdot d^p$ . The Price of Anarchy to network operators under the LI assumption is thus  $PoA^{\#0} = C^\#(\mathbf{F}^0)/C^\#(\mathbf{F}^\#) = (m+1)/m$ , which is maximal for the number  $m = 1$  of links  $\alpha_i$ . The Price of Anarchy to network operators in networks of parallel links is thus upper-bounded by  $PoA_{m=1}^{\#0} = 2$  whereas the Price of Anarchy to end-hosts is unbounded for arbitrary  $p$ .

### 3.3 PI Equilibrium

If the end-hosts  $e_1, \dots, e_K$  are equipped with perfect information, they are in equilibrium if and only if the *selfish* marginal cost of every path to AS  $D$  is the same for every end-host. Under this condition, the cost term  $C^*$  of the PI equilibrium  $\mathbf{F}^+$  to end-hosts can be derived to be  $C^*(\mathbf{F}^+) = d^{p+1} (1 - (p/K)/(p/K + 1)^{(p+1)/p})$  (cf. Appendix A.2). The Price of Anarchy to end-hosts under the PI assumption is

$$PoA^{*+} = \left(1 - \frac{p/K}{(p/K + 1)^{(p+1)/p}}\right) \cdot PoA^{*0} \leq PoA^{*0}.$$

The cost  $C^\#$  of the PI equilibrium  $\mathbf{F}^+$  to network operators is  $C^\#(\mathbf{F}^+) = (m + 1/(p/K + 1)) \cdot d^p$  and the corresponding Price of Anarchy to network operators is

$$PoA^{\#+} = \frac{m + 1/(p/K + 1)}{m} \leq \frac{m + 1}{m} = PoA^{\#0}.$$

Based on the Prices of Anarchy in Table 2, Theorem 1 holds. However, the Prices of Anarchy  $PoA^{*+}$  and  $PoA^{\#+}$  under the PI assumption are dependent on  $K$ , which is the number of end-hosts in the network. If  $K$  is very high, as it is in an Internet context, the Prices of Anarchy under the PI assumption approximate the Prices of Anarchy under the LI assumption. Thus, for scenarios of heterogeneous parallel paths to a destination, the benefit provided by perfect information is undone in an Internet context. In fact, the effect of additional information can even turn negative when considering more general networks, as we will show in the next section.

	LI equilibrium	PI equilibrium
End-host perspective	$\frac{1}{1 - p/(p+1)^{(p+1)/p}}$	$\frac{1 - (p/K)/(p/K+1)^{(p+1)/p}}{1 - p/(p+1)^{(p+1)/p}}$
Network-operator perspective	$\frac{m+1}{m}$	$\frac{m+1/(p/K+1)}{m}$

Table 2: Price of Anarchy for different perspectives and different equilibrium definitions in the network of parallel links (Figure 3).

## 4 The Drawbacks of Information

We will now show that in more general settings, more information for end-hosts can *deteriorate* outcomes of selfish routing. Such a case is given by the general *ladder network* in Figure 4, a natural generalization of the simple topology considered above and a traditional ISP topology [19].

A ladder network of height  $H$  contains  $H$  horizontal links  $h_1, \dots, h_H$ , which represent the rungs of a ladder and have the cost function  $c_{h_i}(f_{h_i}) = f_{h_i}^p$ . Each horizontal link  $h_i$  connects an AS  $A_{i1}$  to AS  $A_{i2}$ , which accommodate the end-hosts  $e_{i1}$  and  $e_{i2}$ , respectively. Every end-host  $e_{i1}$  has the same demand  $d$  towards

the corresponding end-host  $e_{i2}$ . Neighboring rungs of a ladder are connected by vertical links  $v_{ij}$ ,  $i \in \{1, \dots, V = H - 1\}$ ,  $j \in \{1, 2\}$ , where the vertical link  $v_{ij}$  connects the ASes  $A_{ij}$  and  $A_{i+1,j}$  and has the linear cost function  $c_{v_{ij}}(f_{v_{ij}}) = t \cdot f_{v_{ij}}$  with  $t \geq 0$ . We denote a ladder network with this structure and a choice of parameters  $H$ ,  $p$ ,  $d$ , and  $t$  by  $\mathcal{L}(H, p, d, t)$ .

By comparing optima and equilibria, we will prove the following theorem in the following subsections:

**Theorem 2.** *For any ladder network  $\mathcal{L}(H, p, d, t)$ , the Value of Information for both end-hosts and network operators is **negative**, i.e.,  $VoI^* < 0$  and  $VoI^\# < 0$ .*

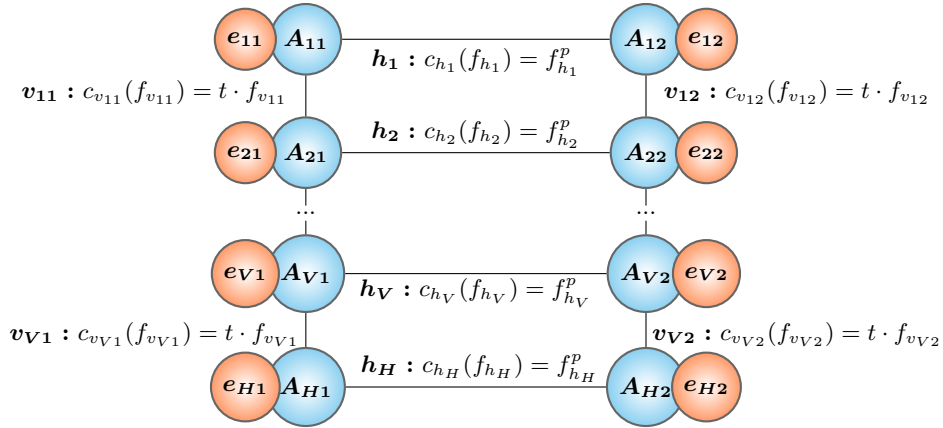


Fig. 4: Example network illustrating the harmful impact of end-host information (Read:  $V = H - 1$ ).

#### 4.1 Social Optima

Both the end-host optimum  $\mathbf{F}^*$  and  $\mathbf{F}^\#$  are equal to the *direct-only* path-flow pattern  $\mathbf{F}^\sim$  that is defined as follows: For every end-host  $e_{i1}$ ,  $F_{(i1,i2),h_i}^\sim = d$  and  $F_{(i1,i2),q}^\sim = 0$  where  $q$  is any other path between  $A_{i1}$  and  $A_{i2}$  than the direct path over link  $h_i$ .

Simple intuition already confirms the optimality of this path-flow pattern. The social cost from the horizontal links is minimized for an equitable distribution of the whole-network demand  $Hd$  onto the  $H$  horizontal links. In contrast, the cost from vertical links  $v_{ij}$  can be minimized to 0 by simply abstaining from using vertical links. In fact, every use of the vertical links is socially wasteful.

More formally, if  $f_{h_i} = d$  for  $i \in \{1, \dots, H\}$  and  $f_{v_{i1}} = f_{v_{i2}} = 0$  for  $i \in \{1, \dots, V\}$ , the marginal costs of the direct path and every indirect path can be easily shown to equal  $(p + 1)d^p$ , given end-host cost function  $C^*$ . Concerning network-operator cost  $C^\#$ , the direct and indirect paths have marginal costs  $p \cdot d^{p-1}$  and  $p \cdot d^{p-1} + 2yt \forall y \in \mathbb{N}_{\geq 1}$ , respectively. The used direct paths thus do not have a higher marginal cost than the unused indirect paths.

## 4.2 LI Equilibrium

Also the LI equilibrium path-flow pattern  $\mathbf{F}^0$  is equal to the direct-only path-flow pattern  $\mathbf{F}^\sim$ . For  $\mathbf{F}^\sim$ , the cost of a direct path  $\pi$  is  $C_\pi(\mathbf{F}^\sim) = F_{(i_1, i_2), \pi}^p = d^p$  and the cost of an indirect path  $\pi'$  is  $f_{h'}^p + \sum_{v \in W_{\pi'}} f_v = d^p + 0 = d^p$ , where  $\pi'$  contains the remote horizontal link  $h'$  and the vertical links  $v \in W_{\pi'}$ . Thus, the LI equilibrium conditions of cost equality are satisfied by  $\mathbf{F}^\sim$ .

As the LI equilibrium is equal to the social optimum both from the end-host perspective and the network-operator perspective, both variants of the Price of Anarchy under the LI assumption are optimal, i.e.,  $PoA^{*0} = PoA^{\#0} = 1$ .

## 4.3 PI Equilibrium

Differently than under the LI assumption, the direct-only flow distribution  $\mathbf{F}^\sim$  is not stable under the PI assumption. An end-host  $e_i$  can improve its individual cost by allocating some traffic to an indirect path  $\pi_k$  (involving the horizontal link  $h_k$ ) and interfering with another end-host  $e_k$ . This reallocation decision will increase the social cost for end-hosts and network operators. In particular, the end-host  $e_k$  that previously used the link  $h_k$  exclusively will see its selfish cost increase. In turn, the harmed end-host  $e_k$  will reallocate some of its traffic to an indirect path in order to reduce its selfish cost  $C_{(e_k)}$ , leading to a process where all end-hosts in the network interfere with each other until they reach a PI equilibrium with a suboptimal social cost for end-hosts and network operators.

Similar to §3.3, we use the condition of marginal selfish cost equality in order to derive the Price of Anarchy under the PI assumption for a ladder network with  $H = 2$ . This derivation, as performed in Appendix A.3, yields the following results for the Price of Anarchy to end-hosts and network operators:

$$PoA_{H=2}^{*+}(p) = 1 + p/12 \quad PoA_{H=2}^{\#+}(p) = 1 + p/3$$

Since the the LI equilibrium is optimal and the PI equilibrium is generally suboptimal on the considered ladder networks, Theorem 2 holds. This finding is confirmed by a case study of the Abilene network (cf. §5), which structurally resembles a ladder topology. The case study also reveals that the negative impact of information is especially pronounced if path diversity is high.

Interestingly, there is an upper bound of the Price of Anarchy to network operators for a *general* ladder network. This bound is given by the following theorem and proven in Appendix A.4:

**Theorem 3.** *For every ladder network  $\mathcal{L}(H, p, d, t)$ , the Price of Anarchy  $PoA^{\#+}$  to network operators is lower than the following upper bound  $PoA_{\max}^{\#+}$ :*

$$PoA^{\#+} \leq PoA_{H, \max}^{\#+} = 1 + \frac{2(H-1)}{3H}p \leq PoA_{\max}^{\#+} = 1 + \frac{2}{3}p$$

## 5 Case Study: Abilene Network

To verify and complement our theoretical insights, we conducted a case study with a real network: we consider the well-known Abilene network, for which topology and workload data is publicly available [14, 15]. We accommodate the Abilene topology into our model as follows. For the demand  $d$  between the 11 points-of-presence, which we consider ASes, we rely on the empirical traffic matrix from the dataset. Concerning the link-cost functions  $c_\ell$ , we model the latency behavior of a link by a function  $c_\ell(f_\ell) = f_\ell^2 + \delta_\ell$ , where  $f_\ell^2$  captures the queuing delay and  $\delta_\ell$  is a constant quantity depending on the geographical distance between the two end-points of link  $\ell$ , approximating the link’s propagation delay.

In order to study the effect of both end-host information and multi-path routing on the Price of Anarchy, we perform the following simulation experiment. First, we compute the social optima  $\mathbf{F}^*$  and  $\mathbf{F}^\#$  for the Abilene network. Second, we simulate the convergence to the Nash equilibria  $\mathbf{F}^0$  and  $\mathbf{F}^+$  for different degrees of multi-path routing, represented by the maximum number of shortest paths that end-hosts consider in their path selection. Once converged, we compute the social cost of the equilibrium traffic distributions and the corresponding Prices of Anarchy.

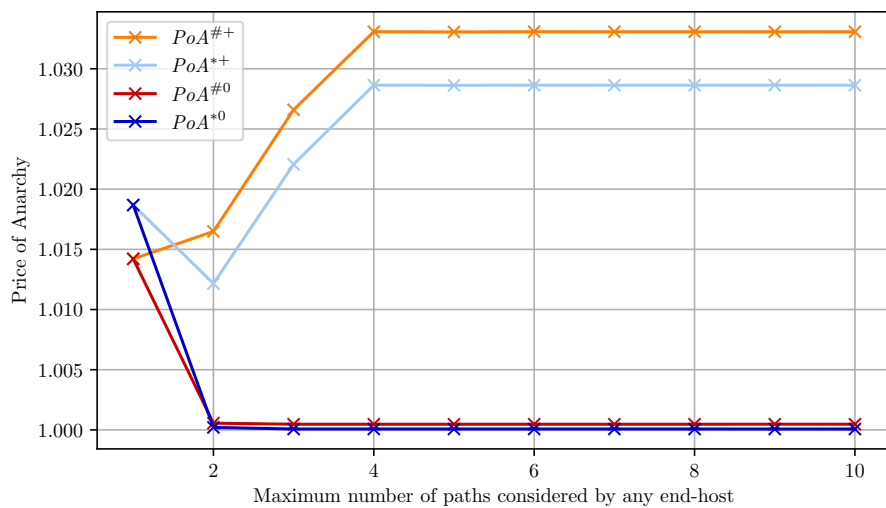


Fig. 5: Abilene network results.

The experiment results in Figure 5 offer multiple interesting insights. Most prominently, if simple shortest-path routing represents the baseline of network-controlled path selection, source-based path selection with latency-only information improves the performance of the network (up to a near-optimum), which confirms findings of prior work [24]. In contrast, path selection with perfect information deteriorates performance, especially for a higher degree of multi-path

routing. Therefore, the potential performance benefits of source-based path selection with multi-path routing are conditional on the amount of information possessed by end-hosts, where a higher degree of information is associated with lower performance. However, while an increasing degree of multi-path routing is associated with worse performance under perfect information, the resulting inefficiency is bounded at a modest level of less than 4 percent for both end-hosts and network operators. The near-optimality of latency-only information in terms of performance and the bounded character of the Price of Anarchy under perfect information reflect the findings from §4 about ladder topologies, which resemble the Abilene topology. Thus, the experiment results not only show that source-based path selection can be a means to improve the performance of a network but also confirm the practical relevance of our theoretical findings.

## 6 Related Work

Inefficiency arising from selfish behavior in networks is well-known to exist in transportation networks and has been thoroughly analyzed with the framework of the Wardrop model [6, 37]. The most salient expressions of this inefficiency is given by the Braess Paradox [4].

Literature on *selfish routing* is often concerned with the discrepancy between optimum and Nash equilibrium: the *Price of Anarchy* [8, 17]. The Price of Anarchy was initially studied for network models (see Nisan et al. [22] for an overview), but literature now covers a wide spectrum, from health care to basketball [29]. Our work has a closer connection to more traditional research questions, such as bounds on the Price of Anarchy for selfish routing. An early result has been obtained by Koutsoupias and Papadimitriou [17], who formulated routing in a network of parallel links as a multi-agent multi-machine scheduling problem.

A different model has been developed by Roughgarden and Tardos [30] who build on the Wardrop model [37] for routing in the context of computer networks. The Price of Anarchy in the proposed routing game is the ratio between the latency experienced by all users in the Wardrop equilibrium and the minimum latency experienced by all users. For different classes of latency functions, the authors derive explicit high bounds on the resulting Price of Anarchy. In a different work, they show that the worst-case Price of Anarchy for a function class can always be revealed by a simple network of parallel links and that the upper bound on the Price of Anarchy depends on the growth rate of the latency functions [27].

The relatively loose upper bounds on the Price of Anarchy of previous works [17, 30] have been qualified by subsequent research. It was found that problem instances with high Prices of Anarchy are usually artificial. By introducing plausible assumptions to make the routing model more realistic, upper bounds on the Price of Anarchy can be reduced substantially. For instance, Friedman [11] shows that the Price of Anarchy is lower than the mentioned worst-case derived by Roughgarden and Tardos [30] if the Nash equilibrium cost is not sensitive to changes in the demand of agents. By computing the Price of Anarchy for a

variety of different latency functions, topologies, and demand vectors, Qiu et al. even show that selfish routing is nearly optimal in many cases [24].

Convergence to Nash equilibria has been studied in the context of congestion games [26] and, in a more abstract form, in the context of potential games [21, 31]. Sandholm [31] showed that selfish player behavior in potential games leads to convergence to the Nash equilibrium and, under some conditions, even to convergence to the social optimum. As the question of equilibrium convergence is traditionally studied separately from the question of equilibrium cost, we do not address convergence issues in this paper.

The study of the effect of incomplete information also has a long tradition [13], but still poses significant challenges [29]. Existing literature in this area primarily focuses on scenarios where players are uncertain about each others' payoffs, studying alternative notions of equilibria such as Bayes-Nash equilibria [35], which also leads to alternative definitions of the price of anarchy such as the Bayes-Nash Price of Anarchy [18, 29] or the price of stochastic anarchy [5]. A common observation of many papers in this area is that less information can lead to significantly worse equilibria [29]. There is also literature on the impact on the Price of Anarchy in scenarios where interacting players only have *local* information, e.g., the evolutionary price of anarchy [33].

However, much less is known today about the role of information in games related to *routing*. In this context, one line of existing literature is concerned with the recentness of latency information. Most prominently, research on the damage done by stale information in load-balancing problems [7, 20] has been applied to routing games by Fischer and Vöcking [10]. This work investigates whether and how rerouting decisions converge onto a Wardrop equilibrium if these rerouting decisions are based on obsolete latency information. Other recent work about the role of information in routing games investigates how the amount of topology information possessed by agents affects the equilibrium cost [1].

Existing work on the subject of source-routing efficiency differs from our work in two important aspects. First, to the best of our knowledge, all existing work on the subject defines the social optimum as the traffic assignment that minimizes the total cost experienced by users, which is indeed a reasonable metric. However, our work additionally investigates the total cost experienced by *links*, i.e., the network operators. Since cost considerations by network operators are a decisive factor in the deployment of source-based path selection architectures, the Price of Anarchy to network operators is an essential metric. Second, although existing work on the topic has investigated the role played by the *recentness* of congestion information or the degree of *topology* information, it does not investigate the role played by the *degree of congestion information* that agents possess. Indeed, a major contribution of our work is to highlight the effects of perfect information, i.e., information that allows agents to perfectly minimize their selfish cost. Latency-only information, which agents are assumed to have in existing work, does not enable agents to perform perfect optimization.

## 7 Conclusion

Motivated by emerging path-aware network architectures, we refined and extended the Wardrop model in order to study the implications of source-based path selection. Our analysis provides several interesting insights with practical relevance. First, the cost of selfish routing to network operators differs from the cost experienced by users. Since network operators are central players in the adoption of path-aware networking, research on the effects of selfish routing thus needs to address the network-operator perspective separately. However, we proved upper bounds on the Price of Anarchy which suggest that selfish routing imposes a low cost on network operators. Second, we found that basic latency information, which can be measured by the end-hosts themselves, leads to near-optimal traffic allocations in many cases. Selfish routing thus causes modest inefficiency even if end-hosts have only imperfect path information and network operators do not disseminate detailed path-load information.

Our model and first results introduce several exciting avenues for future research. First, we note that while we have focused on path-aware network architectures, we hope to apply our model to other practical applications where source routing has currently received much attention, e.g., in the context of segment routing [9] and multi-cast [34]. Furthermore, we aim to obtain a more general understanding of the interactions between the network topology structure and the Price of Anarchy in selfish routing. Moreover, our focus in this paper was on rational players, and it is important to extend our model to account for other behaviors, e.g., players combining altruistic, selfish and Byzantine behaviors.

## 8 Acknowledgements

We gratefully acknowledge support from ETH Zurich, from SNSF for project ESCALATE (200021L\_182005), and from WWTF for project WHATIF (ICT19-045, 2020-2024). Moreover, we thank Markus Legner, Jonghoon Kwon, and Juan A. García-Pardo for helpful discussions that supported and improved this research. Lastly, we thank the anonymous reviewers for their constructive feedback.

## References

1. Acemoglu, D., Makhdoumi, A., Malekian, A., Ozdaglar, A.: Informational braess' paradox: The effect of information on traffic congestion. *Operations Research* **66**(4) (2018)
2. Andersen, D., Balakrishnan, H., Kaashoek, F., Morris, R.: *Resilient overlay networks*, vol. 35 (2001)
3. Barrera, D., Chuat, L., Perrig, A., Reischuk, R.M., Szalachowski, P.: The scion internet architecture. *Communications of the ACM* **60**(6) (2017)
4. Braess, D.: Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung* **12**(1) (1968)



5. Chung, C., Ligett, K., Pruhs, K., Roth, A.: The price of stochastic anarchy. In: International Symposium on Algorithmic Game Theory. pp. 303–314. Springer (2008)
6. Dafermos, S.C., Sparrow, F.T.: The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards B* **73**(2) (1969)
7. Dahlin, M.: Interpreting stale load information. *IEEE Transactions on parallel and distributed systems* **11**(10) (2000)
8. Dubey, P.: Inefficiency of nash equilibria. *Mathematics of Operations Research* **11**(1) (1986)
9. Filsfil, C., Nainar, N.K., Pignataro, C., Cardona, J.C., Francois, P.: The segment routing architecture. In: *IEEE Global Communications Conference (GLOBECOM)* (2015)
10. Fischer, S., Vöcking, B.: Adaptive routing with stale information. *Theoretical Computer Science* **410**(36) (2009)
11. Friedman, E.J.: Genericity and congestion control in selfish routing. In: *43rd IEEE Conference on Decision and Control (CDC)*. vol. 5 (2004)
12. Gupta, A., Vanbever, L., Shahbaz, M., Donovan, S.P., Schlinker, B., Feamster, N., Rexford, J., Shenker, S., Clark, R., Katz-Bassett, E.: Sdx: A software defined internet exchange. *ACM SIGCOMM Computer Communication Review* **44**(4), 551–562 (2015)
13. Harsanyi, J.C.: Games with incomplete information played by “bayesian” players, i–iii part i. the basic model. *Management science* **14**(3) (1967)
14. Knight, S., Nguyen, H.X., Falkner, N., Bowden, R., Roughan, M.: The internet topology zoo. *IEEE Journal on Selected Areas in Communications* (2011)
15. Kolaczyk, E.D.: *Statistical analysis of network data (Datasets)*. Springer (2009), <http://math.bu.edu/people/kolaczyk/datasets.html>
16. Kotronis, V., Kloti, R., Rost, M., Georgopoulos, P., Ager, B., Schmid, S., Dimitropoulos, X.: Stitching inter-domain paths over ixps. In: *Proc. ACM Symposium on SDN Research (SOSR)* (2016)
17. Koutsoupias, E., Papadimitriou, C.: Worst-case equilibria. In: *Annual Symposium on Theoretical Aspects of Computer Science (STACS)* (1999)
18. Leme, R.P., Tardos, E.: Pure and Bayes-Nash price of anarchy for generalized second price auction. In: *IEEE 51st Annual Symposium on Foundations of Computer Science (FOCS)* (2010)
19. Luizelli, M.C., Bays, L.R., Buriol, L.S., Barcellos, M.P., Gaspary, L.P.: Characterizing the impact of network substrate topologies on virtual network embedding. In: *Proceedings of the 9th International Conference on Network and Service Management (CNSM)*. IEEE (2013)
20. Mitzenmacher, M.: How useful is old information? *IEEE Transactions on Parallel and Distributed Systems* **11**(1) (2000)
21. Monderer, D., Shapley, L.S.: *Potential games*. Games and economic behavior (1996)
22. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: *Algorithmic game theory*. Cambridge University Press (2007)
23. Perrig, A., Szalachowski, P., Reischuk, R.M., Chuat, L.: *SCION: A Secure Internet Architecture*. Springer International Publishing AG (2017). <https://doi.org/10.1007/978-3-319-67080-5>, <https://netsec.ethz.ch/publications/papers/SCION-book.pdf>
24. Qiu, L., Yang, Y.R., Zhang, Y., Shenker, S.: On selfish routing in internet-like environments. In: *Proceedings of the ACM SIGCOMM* (2003)
25. Raghavan, B., Snoeren, A.C.: A system for authenticated policy-compliant routing. In: *ACM SIGCOMM Computer Communication Review*. vol. 34 (2004)

26. Rosenthal, R.W.: A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory* **2**(1), 65–67 (1973)
27. Roughgarden, T.: The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences (JCSS)* **67**(2) (2003)
28. Roughgarden, T.: Routing games. *Algorithmic game theory* **18** (2007)
29. Roughgarden, T.: The price of anarchy in games of incomplete information. *ACM Transactions on Economics and Computation (TEAC)* **3**(1) (2015)
30. Roughgarden, T., Tardos, É.: How bad is selfish routing? *Journal of the ACM (JACM)* **49**(2) (2002)
31. Sandholm, W.H.: Potential games with continuous player sets. *Journal of Economic theory* **97**(1), 81–108 (2001)
32. Savage, S., Collins, A., Hoffman, E., Snell, J., Anderson, T.: The end-to-end effects of internet path selection. In: *ACM SIGCOMM Computer Communication Review*. vol. 29 (1999)
33. Schmid, L., Chatterjee, K., Schmid, S.: The evolutionary price of anarchy: Locally bounded agents in a dynamic virus game. In: *Proc. 33rd International Symposium on Distributed Computing (OPODIS)* (2019)
34. Shahbaz, M., Suresh, L., Rexford, J., Feamster, N., Rottenstreich, O., Hira, M.: Elmo: Source-routed multicast for cloud services. *arXiv preprint arXiv:1802.09815* (2018)
35. Singh, S., Soni, V., Wellman, M.: Computing approximate bayes-nash equilibria in tree-games of incomplete information. In: *Proceedings of the 5th ACM conference on Electronic Commerce (EC)* (2004)
36. Trammell, B., Smith, J.P., Perrig, A.: Adding path awareness to the internet architecture. *IEEE Internet Computing* **22**(2), 96–102 (Mar 2018). <https://doi.org/10.1109/MIC.2018.022021673>
37. Wardrop, J.G.: Some theoretical aspects of road traffic research. In: *Inst Civil Engineers Proc London/UK/* (1952)
38. Xu, W., Rexford, J.: Miro: Multi-path Interdomain Routing. In: *Proceedings of the SIGCOMM 2006*

## Appendix A Proofs

### A.1 Parallel Links: End-Host Optimum

In the end-host optimum  $\mathbf{F}^*$ , it holds that for every end-host  $e_k$ ,  $k \in \{1, \dots, K\}$ , the marginal cost of the path over the link  $\beta$  is equal to the marginal cost of any path over a link  $\alpha_i$ ,  $i \in \{1, \dots, m\}$ . Using this insight, the end-host optimum can be derived by the solution of the following equation:

$$\frac{\partial}{\partial F_{(k,D),\beta}} C^* = \frac{\partial}{\partial F_{(k,D),\alpha_i}} C^*$$

As it holds that

$$\frac{\partial}{\partial F_{(k,D),\ell}} C^* = \frac{\partial}{\partial f_\ell} C^*$$

the optimal link-flow pattern  $\mathbf{f}^*$  is obtained by solving:

$$\frac{\partial}{\partial f_\beta} (f_\beta \cdot f_\beta^p) = \frac{\partial}{\partial f_{\alpha_i}} (f_{\alpha_i} \cdot d^p),$$

which yields  $f_\beta^* = d/\sqrt[p]{p+1}$ .

Any path-flow pattern  $\mathbf{F}^*$  that produces a link-flow pattern with  $f_\beta = f_\beta^*$  and  $\sum_{\alpha_i} f_{\alpha_i} = d - f_\beta^*$  is thus optimal from the perspective of end-hosts. The total cost of such an optimal path-flow pattern  $\mathbf{F}^*$  is given by

$$C^*(\mathbf{F}^*) = f_\beta^* \cdot f_\beta^{*p} + (d - f_\beta^*) \cdot d^p = d^{p+1} \left(1 - p/(p+1)^{(p+1)/p}\right)$$

## A.2 Parallel Links: PI Equilibrium

In the network from Figure 3, the selfish cost function of an end-host  $e_k$  can be simplified to

$$C_{(e_k)}^*(\mathbf{F}) = F_{(k,D),\beta} \cdot f_\beta^p + \sum_{\alpha_i} F_{(k,D),\alpha_i} \cdot d^p$$

Since it holds that  $\partial/\partial F_{(k,D),\beta} f_\beta = 1$ , the marginal selfish costs for the paths over link  $\beta$  and  $\alpha_i$  are given by

$$\begin{aligned} \partial/\partial F_{(k,D),\beta} C_{(e_k)}^*(\mathbf{F}) &= f_\beta^p + F_{(k,D),\beta} \cdot p \cdot f_\beta^{p-1} \\ \partial/\partial F_{(k,D),\alpha_i} C_{(e_k)}^*(\mathbf{F}) &= d^p. \end{aligned}$$

An equilibrium under the PI assumption is characterized by the equality of these selfish marginal costs.

Note that for every end-host  $e_k$ , the marginal selfish cost of every path over a link  $\alpha_i$  is the same, namely  $d^p$ . By marginal selfish cost equality, the marginal selfish costs  $\partial/\partial F_{(k,D),\beta} (C_{(e_k)}^*)$  must be equal to  $d^p$  for every end-host  $e_k$  and thus also equal across all end-hosts. This condition is only satisfied if every end-host  $e_k$  has the same flow on the path over link  $\beta$ , i.e.,  $F_{(m,D),\beta} = F_{(n,D),\beta}$  for all end-hosts  $e_m, e_n$ . As  $f_\beta = \sum_{e_k} F_{(k,D),\beta}$ , the fact  $\sum_{e_k} F_{(k,D),\beta} = K \cdot F_{(k,D),\beta}$  implies  $F_{(k,D),\beta} = f_\beta/K$  for every end-host  $e_k$ .

This knowledge about  $F_{(k,D),\beta}$  allows to simplify the selfish marginal cost equation and to compute the PI equilibrium. By inserting  $f_\beta/K$  for  $F_{(k,D),\beta}$ , the selfish marginal cost equality reads

$$f_\beta^p + f_\beta/K \cdot p \cdot f_\beta^{p-1} = d^p.$$

The solution of this equation yields the PI equilibrium link flow  $f_\beta^+ = d/\sqrt[p]{p/K+1}$ . The PI equilibrium  $\mathbf{F}^+$  is thus given by every path-flow pattern that satisfies the following conditions for every end-host  $e_k$ :

$$F_{(k,D),\beta} = f_\beta^+/K \quad \wedge \quad \sum_{\alpha_i} F_{(k,D),\alpha_i} = (d - f_\beta^+)/K.$$

The cost term  $C^*$  of the PI equilibrium  $\mathbf{F}^+$  to end-hosts is thus

$$C^*(\mathbf{F}^+) = d^{p+1} \left(1 - (p/K)/(p/K+1)^{(p+1)/p}\right).$$

### A.3 Ladder Network: PI Equilibrium

In order to compute the Price of Anarchy under the PI assumption for a general ladder network, we start by computing the Price of Anarchy for the simple ladder network of  $H = 2$ . Conforming to the PI equilibrium conditions in §2.4, the PI equilibrium is given by the solution of the following equilibrium equation system  $\mathcal{E}_2$  that formalizes the marginal cost equality:

$$\begin{cases} f_{h_1}^p + F_1' p f_{h_1}^{p-1} = f_{h_2}^p + t(f_{v_{11}} + f_{v_{12}}) + F_1(p \cdot f_{h_2} + 2t) \\ f_{h_2}^p + F_2' p f_{h_2}^{p-1} = f_{h_1}^p + t(f_{v_{11}} + f_{v_{12}}) + F_2(p \cdot f_{h_1} + 2t) \end{cases}$$

where  $F_1' = F_{(11,12),h_1}$  and  $F_1 = F_{(11,12),v_{11}h_2v_{12}}$  and the analogous abbreviations have been made by  $F_2'$  and  $F_2$  for the direct and indirect flow of end-host  $e_{21}$ . Due to demand constraints, it holds that  $F_1' = d - F_1$  and  $F_2' = d - F_2$ . Due to the symmetry of the equation system, it is possible to conclude that  $F_1' = F_2'$  and  $F_1 = F_2$ . Since  $f_{h_1} = F_1' + F_2$  and  $f_{h_2} = F_2' + F_1$ , using the symmetry yields  $f_{h_1} = f_{h_2} = d - F_1 + F_1 = d$ . Furthermore, the flow on the vertical links  $v_{11}$  and  $v_{12}$  can be expressed as follows:  $f_{v_{11}} = f_{v_{12}} = F_1 + F_2 = 2F_1$ . The equilibrium equation system can thus be reduced to the single equation:

$$\begin{aligned} d^p + (d - F_1) \cdot \lambda &= d^p + 2t \cdot (2F_1) + F_1 \cdot (\lambda + 2t) \\ \iff \lambda d - (6t + 2\lambda)F_1 &= 0 \iff F_1 = \lambda d / (6t + 2\lambda) \end{aligned}$$

where  $\lambda = p \cdot d^{p-1}$ .

Based on this solution for the path flow  $F_1 = F_{(11,12),v_{11}h_2v_{12}}$ , all other path flows can be derived, which yields the following terms for the two perspectives on the Price of Anarchy:

$$\begin{aligned} PoA_{H=2}^{*+}(d, t, p) &= (2d^{p+1} + 2t(2\lambda d / (6t + 2\lambda))^2) / (2 \cdot d^{p+1}) \\ PoA_{H=2}^{\#+}(d, t, p) &= (2d^p + 4t\lambda d / (6t + 2\lambda)) / (2 \cdot d^p) \end{aligned}$$

The Price of Anarchy for all ladder networks with  $H = 2$  is obtained by computing an upper bound on the Price of Anarchy in terms of demand  $d$  and parameter  $t$ :

$$PoA_{H=2}^{*+}(p) = 1 + p/12 \quad PoA_{H=2,\max}^{\#+}(p) = 1 + p/3$$

### A.4 Ladder Network: Proof of Theorem 3

We start by observing that  $PoA_{H,\max}^{\#+}$  is given by the limit in  $t$  and is only dependent on the flows on vertical links  $v$ :

$$PoA_{H,\max}^{\#+} = 1 + \lim_{t \rightarrow \infty} (t \cdot \sum_v f_v) / (H \cdot d^p),$$

where we used that  $\lim_{t \rightarrow \infty} \sum_h f_h^p = H \cdot d^p$  as vertical links become infinitely expensive. We need only characterize the sum of vertical link flows  $f_V = \sum_v f_v$ ,

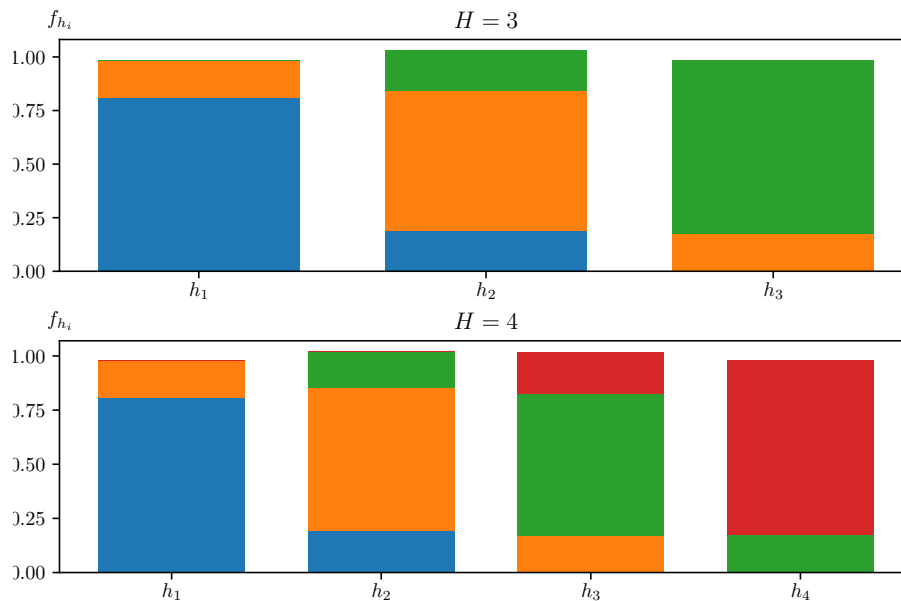


Fig. 6: Traffic distribution over horizontal links of a ladder network in PI equilibrium  $\mathbf{F}^+$  (for  $p = 2$ ,  $t = 1$ ,  $H = 3, 4$ ).

for which we use an argument based on the structure of the equilibrium equation system  $\mathcal{E}_H$ .

For setting up  $\mathcal{E}_H$ , we consider Figure 6 which illustrates numerically computed equilibria for some  $H > 2$ . The figure shows the traffic distribution on the horizontal links of a ladder network, where different-color flow shares correspond to flows of different end-hosts. Figure 6 shows two insights that are relevant for setting up  $\mathcal{E}_H$ . First, the path-flow pattern  $\mathbf{F}^+$  only contains non-zero flows on paths that deviate at most one ladder level from the originating end-host (for high enough  $t$ ). Second, the path-flow pattern is symmetric with respect to the horizontal axis of the ladder network. The variables in  $\mathcal{E}_H$  can thus be assigned to the indirect path flows as displayed in Figure 7. Variable assignments for higher  $H$  work analogously to Figure 7a (for odd  $H$ ) and Figure 7b (for even  $H$ ).

With these variables and the knowledge about the equilibrium traffic distribution, the equation system  $\mathcal{E}_H$  can be set up for all values for  $H$ . Table 3 lists the equation systems  $\mathcal{E}_H$  for all  $H$ . All equations in a system  $\mathcal{E}_H$  are of the form  $E_{a,b}(F_c, F_d, F_f) = 0$ , where

$$E_{a,b}(F_c, F_d, F_f) = \lambda d - (at + b\lambda) \cdot F_c - 2t \cdot F_d - \lambda \cdot F_f.$$

The set  $E(\mathcal{E}_H)$  contains all left-hand side terms  $E_{a,b}$  of the equations in  $\mathcal{E}_H$ . Let the sum  $\Sigma(\mathcal{E}_H)$  of an equation system  $\mathcal{E}_H$  be the equation  $\sum_{E \in \mathcal{E}_H} E_{a,b} = 0$ . It

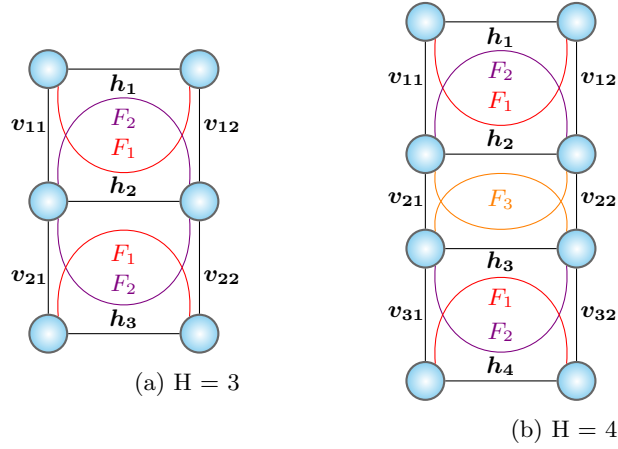


Fig. 7: Variable assignments to indirect path flows in a ladder network for computing the PI equilibrium.

holds that for all  $H$ ,  $\Sigma(\mathcal{E}_H)$  is the equation

$$(H - 1)\lambda d - (6t + 2\lambda) \cdot F_1 - (6t + 3\lambda) \sum_{2 \leq u \leq H-1} F_u = 0.$$

Solving equation  $\Sigma(\mathcal{E}_H)$  for  $F_V = \sum_{1 \leq u \leq H-1} F_u$ , we obtain

$$F_V = \frac{(H - 1)\lambda d}{\rho(t) \cdot (6t + 3\lambda)}$$

where  $\lim_{t \rightarrow \infty} \rho(t) = 1$ .

Due to the horizontal and vertical symmetry of the PI equilibrium on the ladder network, it holds that  $f_V = 4 \cdot F_V$ . Inserting  $f_V$  into  $PoA_{H, \max}^{\# +}$  yields

$$1 + \lim_{t \rightarrow \infty} \frac{t}{H \cdot d^p} \cdot \frac{4(H - 1)\lambda d}{\rho(t)(6t + 3\lambda)} = 1 + \frac{2(H - 1)}{3H} \cdot p.$$

Taking the limit of this term for  $H \rightarrow \infty$  results in

$$PoA_{\max}^{\# +} = 1 + 2/3 \cdot p.$$

$H$	$\mathcal{E}_H$
$H = 2$	$\left\{ E_{6,2}(F_1, 0, 0) = 0 \right.$
$H = 3$	$\left. \begin{array}{l} E_{4,2}(F_1, F_2, 0) = 0 \\ E_{4,3}(F_2, F_1, 0) = 0 \end{array} \right\}$
Even $H \geq 4$ Odd $j, 3 \leq j \leq H - 3$	$\left\{ \begin{array}{l} E_{4,2}(F_1, F_2, 0) = 0 \\ E_{4,2}(F_2, F_1, F_3) = 0 \\ E_{4,2}(F_j, F_{j+1}, F_{j-1}) = 0 \\ E_{4,2}(F_{j+1}, F_j, F_{j+2}) = 0 \\ E_{6,2}(F_{H-1}, 0, F_{H-2}) = 0 \end{array} \right.$
Odd $H \geq 5$ Odd $j, 3 \leq j \leq H - 4$	$\left\{ \begin{array}{l} E_{4,2}(F_1, F_2, 0) = 0 \\ E_{4,2}(F_2, F_1, F_3) = 0 \\ E_{4,2}(F_j, F_{j+1}, F_{j-1}) = 0 \\ E_{4,2}(F_{j+1}, F_j, F_{j+2}) = 0 \\ E_{4,2}(F_{H-2}, F_{H-1}, F_{H-3}) = 0 \\ E_{4,3}(F_{H-1}, F_{H-2}, 0) = 0 \end{array} \right.$

 Table 3: Equation systems  $\mathcal{E}_H$  characterizing PI equilibrium for all  $H$ .