Selfish Routing in the Internet: Modeling TCP-like Flows

Aditya Akella Shuchi Chawla Srinivasan Seshan

Computer Science Department
Carnegie Mellon University
{aditya,shuchi,srini+}@cs.cmu.edu

1 Introduction

The Internet of today is a large collection of independently administered autonomous entities. These entities typically take unilateral decisions, such as selecting a path to route their packets, in order to maximize their own utility. In this note, we study the effect of selfish routing on the efficient performance of the Internet.

Several studies in the past have cast this question as a game between selfish agents and have tried to quantify what is called the price of anarchy [6]—the ratio of performance of the network at the Nash Equilibrium of the game, to the optimal performance. In a series of papers, Roughgarden et al study the cost of Nash equilibrium, when selfish flows in a network pick routes so as to minimize their latencies (see for example, [9, 7]). They show that the cost of a Nash equilibrium can, in general, be arbitrarily larger than the cost of the optimal routing, except when the latencies on links in the network depend linearly on the total flow on them. The problem has also been studied by Koutsoupias and Papadimitriou [4], and by Mavronicolas et al [5] and Czumas et al [3], but on a slightly different model. Koutsoupias et al give bounds on the price of anarchy when agents own unit amount of flow and pick paths probabilistically.

While seminal in their contributions, the aforementioned studies have a common drawback: they are too simplistic in their modeling of the behavior of flows on the Internet. In this paper, we study the price of anarchy in a more realistic model of the Internet. We assume that the agents are *unsplittable TCP flows*. TCP is the predominant model of transport in today's Internet used by over 80% of the bytes carried by the network. TCP uses a congestion control algorithm to help adapt its rate of flow to the bandwidth available on its path. In this setting, we assume that selfish users (TCP flows) aim to minimize the time that it takes to complete a file transfer (time between transmitting the first and the last byte on the flow), and pick the route that minimizes this time. We show that choosing a route that minimizes completion time is equivalent to choosing one that maximizes the bandwidth available to the flow. Thus, each flow picks a path that has the maximum bottleneck bandwidth available for the flow. The bandwidth available on a link is divided among all flows using that link, in inverse proportion to their end-to-end latencies (or round-trip times), since TCP is RTT-fair [11]. Note that we do not consider the slow start phase of TCP (we do not consider "short-lived" flows). We assume that the system quickly converges to a stable state, and evaluate the performance in this stable state. We also do not consider the dynamic arrival and departure of flows.

2 The model and some observations

We are given a network with capacities, or bandwidths b_e, on edges. There are n agents — agent i is a flow from node s_i to t_i. Each agent picks a path P_i to route its flow. Let the transmission delay, or round trip time (RTT), for routing packets on path P_i be r_i. The RTT of a link depends on the physical properties of the link and its load. Since we are considering TCP senders, that adapt the rate of flow to the bandwidth of the link (*i.e.* flows are *elastic*), the load on each link stays constant on average; thus the RTT of the path, r_i, stays independent of the number of flows on any link. Next, we describe how we model elastic TCP flows.

After all agents have chosen their paths, for any edge e, let F_e be the set of agents whose paths contain the edge e. The bandwidth *available* to agent i on edge e is given by $\frac{1}{Z_e}\frac{b_e}{r_i}$, where $Z_e = \sum_{i\in F_e}\frac{1}{r_i}$. That is, the total bandwidth available on edge e is divided among the flows using this edge in inverse proportion of their RTTs. The bottleneck bandwidth *used* by flow i, B_i, is the minimum bandwidth available to it on edges in its path. The time taken by i to complete its file transfer is given by $\frac{S_i}{B_i}+r_i$, where S_i is the size of the file that it is trying to transfer. Each agent aims to minimize the file transfer time. Assuming that the transmission delay is a negligible component of this time, this is equivalent to maximizing the bandwidth B_i. The performance of the network is evaluated by the total amount of bandwidth used by all flows.

We observe that a pure Nash equilibrium for this game always exists, although it may not be unique. The existence can be established by the following simple argument. Start with any set of paths P_i for each flow i. As long as there is a flow i that is unhappy (that is, it can improve the bandwidth available to it by switching to a different path $\hat{P}_i$), change the path of flow i from P_i to $\hat{P}_i$. Let the bandwidth available to flow i at a particular step be a_i. Arrange the bandwidths a_i in decreasing order to obtain the sequence $\hat{a}_i$. Then consider the function $\sum_i T^i \hat{a}_i + \sum_i a_i$, where T is the sum of bandwidths of all the edges in the network. Some simple manipulation shows that at every step the value of this function increases strictly, but cannot exceed a fixed bound. Hence, the process terminates and gives a Nash equilibrium. We defer the proof to the full version of this paper [1].

Lemma 1 *The selfish routing game always has a pure Nash equilibrium, however it may not be unique.*

As an example, consider the network in Figure 1(a). All links have bandwidth 1. There are $\frac{n}{2} - 1$ flows from s_0 to t_0 and one flow each from s_i to t_i, $i \geq 1$. The transmission delay of each link is set so that all flows have equal RTTs. One Nash equilbrium in this game is defined by all flows from s_0 to t_0 going via all the x_i. In this equilibrium, each flow gets $\frac{2}{n}$ bandwidth, and the total bandwidth used is 2. In another equilibrium, which is also the optimal solution for this network, the flows from s_0 to t_0 take the path $s_0 x_1 x_2 t_0$. In this equilibrium, the flows from s_0 to t_0 and the flow from s_1 to t_1 get bandwidth $\frac{2}{n}$, whereas, the rest of the flows get bandwidth 1. Thus the total bandwidth used is $\frac{n}{2} + 1$, which is $\Omega(n)$ times that used in the previous case.

3 Results and Discussion

3.1 The price of anarchy

In this section we explore the price of anarchy under the above model, for different kinds of networks. First note that the above example shows that in the worst case in a general network, the performance of a Nash equilibrium can be arbitrarily worse than that of the optimal solution. Thus we get the following:

Lemma 2 *The price of anarchy in a network with n flows can be as large as $\Omega(n)$.*

Note that the example above uses a network that has a diameter of $\Omega(n)$. This motivates us to consider the case of a network where all flows use short paths. The price of anarchy in this case improves considerably. This is encouraging as the diameter of the Internet is known to be $O(\log v)$, where v is the number of nodes. Also, it has been shown that typical path length in the Internet is $O(\log v)$.

Lemma 3 *The price of anarchy in a network where all agents use paths of length at most D, is at most $D+1$.*

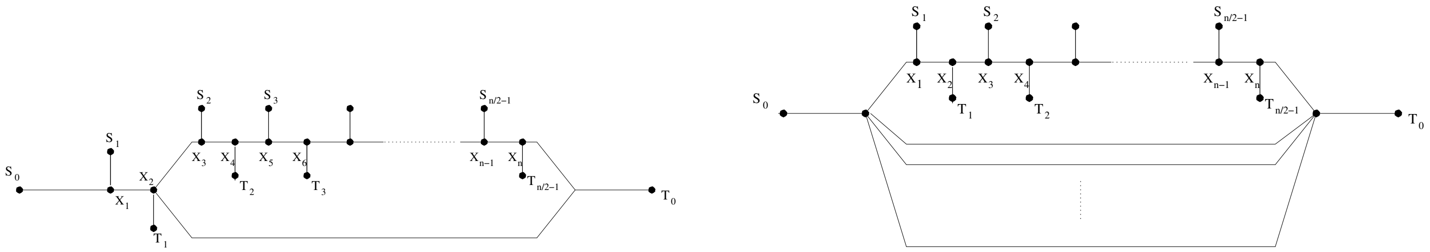
Proof: Let $P_1, \cdots, P_n$ be the paths used by flows in Nash equilibrium, with bottleneck bandwidths $b_1, \cdots, b_n$ respectively, and let $Q_1, \cdots, Q_n$ be the corresponding paths in the optimal solution. Consider the solutions $S_i = Q_1, \cdots, Q_i, P_{i+1}, \cdots, P_n$, $i = 0 \cdots n$. Let T_i be the total bandwidth used by flows in S_i. We use a function, "buffer-bandwidth" W_i, such that $W_0 = 0$ and $\forall i$, $W_i \geq 0$. We show that $T_i + W_i \leq T_{i-1} + W_{i-1} + Db_i$. Summing over all i, we get $T_n \leq T_0 + D\sum_i b_i = (D+1)T_0$, which implies the result.

Note that the only difference between S_{i-1} and S_i is that flow i shifts from P_i to Q_i. This frees b_i amount of bandwidth on each link in the path P_i. In the worst case, this bandwidth becomes available to different flows on every link. Since there are at most D such links, the increase in bandwidth is at most Db_i. If this bandwidth is used by other flows, it leads to an increase in T_i. The rest of the unused bandwidth is buffered, or added to W_{i-1} to obtain W_i. Now consider the new bandwidth available to the flow i. Let e be the bottleneck link on path Q_i. If e contains other flows, there is no increase in bandwidth. If it doesn't contain any flows, then either it did not contain any other flows in S_0, in which case, its bandwidth is at most b_i. Otherwise, it contained flows in S_0, but all these flows moved to different paths. In the latter case, the bandwidth of link e was buffered in W_i. Accordingly, we decrement W_i by the amount of increase in T_i. This implies that $T_i + W_i \leq T_{i-1} + W_{i-1} + Db_i$, and we are done. ■

Next we observe that although one Nash equilibrium in Figure 1(a) is arbitrarily bad with respect to the optimal solution, there is another equilibrium, that is in fact optimal. This leads us to conjecture that in a network with bounded bandwidth, there always exists a Nash equilibrium that has a low price of anarchy. Below we give some evidence for this conjecture.

Conjecture 1 *Let the ratio between the maximum bandwidth of any link and the minimum bandwidth of any link in the network be B. Then the optimistic price of anarchy is at most $O(B)$.*

Assume without loss of generality, that the minimum bandwidth in the network is 1, and the maximum is B. For a path P, let $p_1 \leq p_2 \leq \cdots$ denote the bandwidths



(a) Bounded bandwidth ratio

(b) Unbounded bandwidth ratio

Figure 1: Lower bound on the price of anarchy in general networks

available to a flow along links in this path. We say that a path P dominates path P' if there exists i, such that, $p_j = p'_j$ for all $j < i$, and $p_i > p'_i$. Suppose that instead of just maximizing their bottleneck bandwidth, agents were trying to achieve a stronger goal—if there are several paths with the same bottleneck bandwidth, then among these, pick the path that dominates all the other available paths. If a Nash equilibrium exists in this model, we show that it has a low cost.

Lemma 4 *Let the bandwidth ratio be bounded by B. Consider a Nash equilibrium in which each agent picks a path that dominates all other paths available to him. The price of anarchy for this equilibrium is at most $O(B)$.*

Proof Sketch: If no player in this equilibrium has a bottleneck bandwidth of less than 1, then we are done, because no player can obtain a bandwidth of more than B. Suppose there are players with bottleneck bandwidth less than 1. Let F be the set of such flows. We will show that it is not possible for these flows to collectively increase the total bandwidth occupied by them. Let E be the set of all links that are bottleneck links for flows in F (there can be multiple such links for some flows). Then, E forms a multi-commodity cut that is fully saturated. Flows that only utilize one link in this set clearly cannot increase their bandwidth without harming other flows. Consider flows that use more than one edge in E. Then if we remove all but any one of the links used by these flows, the remaining set still forms a saturated cut with respect to the flows using this set (because the flow picked a dominating path)[1]. Thus again, the flows using this set cannot collectively increase their bandwidth. ■

Note that when the bandwidth ratio is unbounded, even the optimistic price of anarchy can be quite high, as can be seen from the simple extension of the above example given in Figure 1(b). In this example, the links through all the x_i have bandwidth $n/2$ each, whereas the $n/2-1$ bottom links have bandwidth $1-\epsilon$ each.

[1]Compare this to the cut defined by the bad Nash equilibrium in Figure 1(a).

3.2 The benefit from using limited central control

The above results are quite discouraging, as in each case, the price of anarchy is quite high. In this section, we examine the improvement in performance of the network, when a central authority can control a fraction of the flows and route them optimally, while the rest of the flows are still selfish. Such a model was first considered by Roughgarden in [8].

First observe from example 1(a) that even if a central authority controls half of the flows, specifically those between s_i and t_i for $i \geq 1$, the price of anarchy remains as before. Thus it seems unlikely that central control would alleviate the problem in this game. In fact, we can construct a single-source single-sink network in which, even if the central authority controls a $(1-\epsilon)$ fraction of the flows, for a constant ϵ, the price of anarchy is $O(n)$.[2] Instead we consider a simpler network model in which each flow has the same source and sink and has a choice between several parallel links. Note that this is exactly the same as the model considered by Roughgarden [8], except that our objective function and latency function is different from the one used by them.

Lemma 5 *In the single source and single sink model, with parallel links, the price of anarchy is at most 2. Further if a central authority controls an α fraction of the flows, then the price of anarchy is at most $2-\alpha$.*

Proof: Without loss of generality, we can assume that there are at most n links, with bandwidths $b_1 \geq b_2 \geq \cdots \geq b_n$, because both Nash equilibrium and the optimal solution use only the n highest bandwidth ones. The optimal solution simply puts one flow on each link and obtains a bandwidth of $\sum_i b_i$. Now consider a Nash equilibrium. The bandwidth of any link not used in this

[2]See the full paper [1] for details.

equilibrium is at most the minimum bandwidth used by any flow. In the worst case, all unused links not used have the same bandwidth, and all flows occupy the same bandwidth. Thus, if there are l unused links, at most $n/(n+l)$ fraction of the total bandwidth is used by Nash. The worst case is achieved when $l = n - 1$, and in this case, the ratio is at most 2. Now, the strategy for the central authority is to compute a Nash equilibrium for the $(1-\alpha)n$ flows, and route the rest of the αn flows on αn highest bandwidth links, thats are not used in the Nash. As before, in the worst case, the ratio is at most $(n - \alpha n + l)/n$, which is at most $2 - \alpha$. ■

4 Future work and open problems

Our work initiates the study of the price of anarchy in a realistic model of the Internet routing game. In a similar effort, Qiu et al [10] show experimentally, that the price of anarchy in networks representing real ISP backbone topologies is quite low. We believe that realistic selfish routing would span many ISPs and would result in a different results than that shown in [10]. Currently, we are also investigating the price of anarchy on special topologies like large power-law graphs and highly connected graphs. Another interesting open problem is to design a mechanism which regulates selfish flows and achieves near optimal performance. Finally, an interesting model is to allow end users to only pick the first hop of their path to the destination. This corresponds to multihomed end-networks, a widely employed strategy to improve performance in the Internet [2]. We expect the performance of the network to improve in this situation, although Figure 1(b) shows that we may not see any improvement in an arbitrary worst case network.

Acknowledgments

We thank Luis von Ahn for helpful discussions.

References

[1] A. Akella and S. Chawla. *Realistic models for selfish routing in the Internet.* http://www.cs.cmu.edu/~shuchi/papers/selfishrouting.ps.

[2] A. Akella, B. Maggs, S. Seshan, A. Shaikh and R. Sitaraman. *A measurement-based analysis of multihoming.* ACM SIGCOMM 2003.

[3] A. Czumaj and B. Vöcking. *Tight bounds for worst-case equilibria.* In SODA, 2002, 413–420.

[4] E. Koutsoupias and C.H. Papadimitriou. *Worst-case equilibria.* In STACS, 1999, 404–413.

[5] M. Mavronicolas and P. Spirakis. *The price of selfish routing.* In STOC, 2001, 510–519.

[6] C.H. Papadimitriou. *Algorithms, Games and the Internet.* In STOC, 2001, 749–753.

[7] T. Roughgarden. *Selfish Routing.* PhD. Thesis, Cornell university, May 2002.

[8] T. Roughgarden. *Stackelberg scheduling strategies.* In STOC, 2001, 104–113.

[9] T. Roughgarden and E. Tardos. *How bad is selfish routing?* In JACM, 49(2):236–259, 2002.

[10] L. Qiu, Y. Yang, Y. Zhang and S. Shenker. *On Selfish Routing in Internet-Like Environments.* ACM SIGCOMM 2003.

[11] M. Allman and V. Paxson. *TCP congestion control.* RFC 2581, Internet Engineering task Force, April 1999.