How Travel Demand Affects the Cost of Selfish Routing

Steven O'Hare, Richard Connors & David Watling: Institute for Transport Studies, University of Leeds

The cost of selfish routing refers to the performance loss that occurs when travellers in a road network select routes to minimise their individual travel costs, thereby following the User Equilibrium (UE) routing principle, rather than conforming to System Optimal routing (SO) that minimises the total network travel cost. The aim of this research is to examine how travel demand gives rise to variations in this efficiency loss. This work has potential future implications for policies such as road pricing that seek to push a UE assignment towards SO, as it provides understanding of how demand affects the gap between UE and SO and therefore highlights where the greatest benefits from such schemes can be derived. In this paper we identify limitations in the standard measure of the efficiency loss, the Price of Anarchy, and propose an alternative that addresses them: the Price of Anarchy Delays ($PoAD$). We then show analytically, using parallel link networks with linear cost functions, and numerically, for a general network in the form of Sioux Falls, that, as demand increases, the variation of $PoAD$ is dependent upon the differing nature of the *activation* of links; i.e. when the flow on a link first becomes non-zero, under the UE and SO models.

To date the standard measure that has been used to quantify the inefficiency of selfish routing is the 'Price of Anarchy' (PoA), which is defined as the ratio of the Total Travel Cost (TTC) under UE routing to the TTC under SO routing. For road networks with polynomial form cost functions Roughgarden and Tardos (2002: Journal of the Acm, 49, 236-259) and Roughgarden (2003: Journal of Computer and System Sciences, 67, 341-364) have shown that this measure has tight upper bounds, which are independent of the demand structure and network topology, and that depend only on the highest power of the cost functions in a given network. More recent numerical studies by Wu et al. (2008: Journal of Transportation Systems Engineering and Information Technology, 8, 69-74) and Youn et al. (2008: Physical Review Letters, 101) have shown that PoA varies with supply and demand structure. However, no explanations are provided for the patterns uncovered, and all but one of the examples they include utilise network topologies and travel demand structures that are not realistic for road networks; for example both studies use random and scale-free network topologies, which are typically non-planar and so contradict the near planarity that is known to exist in real road network topologies (Barthelemy, 2011: Physics Reports-Review Section of Physics Letters, 499, 1-101). A plausible example is provided by Youn et al. (2008); see Figure 1, illustrating three differing patterns of PoA varying with travel demand in single OD sub-networks of three different cities with no accompanying explanation. In this work we attempt to understand the mechanisms underlying the variation of the cost of selfish routing with travel demand that give rise to patterns such as those shown in Figure 1.

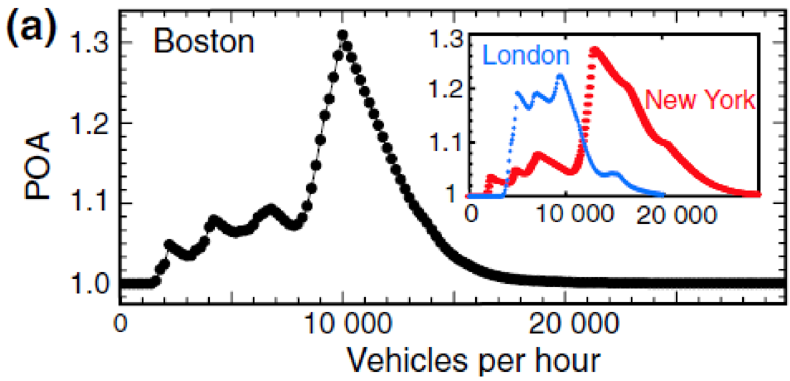


Figure 1: PoA against Demand from Youn et al (2008)

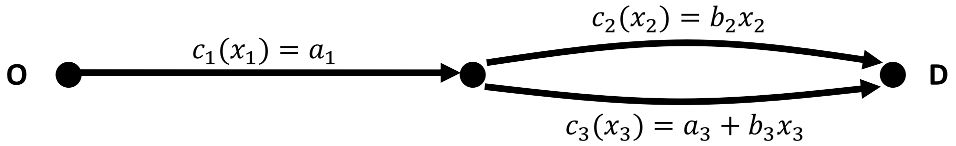
We begin by noting that the maximum PoA in Figure 1 of approximately 1.3 is significantly lower than the upper bound; derived using the results of Roughgarden (2003), of approximately 3.5 for such networks. The formulation of TTC reveals why the PoA upper bound is typically not achieved for real networks. Thus, assuming a single OD pair serving a demand q with K routes to simplify notation and supposing that wlog routes $1, 2, \dots, K$ are ordered such that route costs $C_1(0) \le C_2(0) \le \cdots \le C_K(0)$, we have:

$$TTC = \sum_{k=1}^{K} f_k C_k(f_k) = \sum_{k=1}^{K} f_k[C_k(0) + C_k(f_k) - C_k(0)] = \sum_{k=1}^{K} f_k C_k(0) + \sum_{k=1}^{K} f_k[C_k(f_k) - C_k(0)]$$

Now as $C_1(0) \le C_2(0) \le \cdots \le C_K(0)$ we can define $\gamma_k = C_k(0) - C_1(0)$ for $k = 1, 2, 3, \dots, K$ to represent the additional free-flow costs of the longer routes $k = 2, 3, \dots, K$, such that $C_1(0) \le C_1(0) + \gamma_2 \le \cdots \le C_1(0) + \gamma_k$. Substituting the γ_k into the first term of TTC:

$$TTC = \sum_{k=1}^{K} f_k[C_1(0) + \gamma_k] + \sum_{k=1}^{K} f_k[C_k(f_k) - C_k(0)] = \sum_{k=1}^{K} f_k C_1(0) + \sum_{k=2}^{K} f_k \gamma_k + \sum_{k=1}^{K} f_k[C_k(f_k) - C_k(0)]$$
$$= C_1(0)q + \sum_{k=2}^{K} f_k \gamma_k + \sum_{k=1}^{K} f_k[C_k(f_k) - C_k(0)] \quad (1)$$

This shows that TTC can be decomposed into a sum of: 1) the free-flow travel cost of routing all demand by the shortest path, 2) the additional free-flow travel costs incurred by those flows forced to use longer routes and 3) the travel delays due to congestion on all routes. Note that all three of these cost components appear in the PoA ratio and that the first free-flow component is independent of the routing strategy; i.e. it takes the same value for UE and SO. To illustrate the influence of the free-flow component on the PoA, consider the following single OD network (where a_1, a_3, b_2, b_3 are positive coefficients) serving a demand q from **O** to **D**:



Using equation (1), with flows $x_2 = x$ and $x_3 = q - x$, the PoA in this network is:

$$PoA = \frac{TTC^{UE}}{TTC^{SO}} = \frac{a_1 q + [a_3(q - x_{UE})^2] + [b_2 x_{UE}^2 + b_3(q - x_{UE})^2]}{a_1 q + [a_3(q - x_{SO})^2] + [b_2 x_{SO}^2 + b_3(q - x_{SO})^2]} \quad (2)$$

All demand q uses link 1, and the link flows x and $q - x$ are independent of the value of a_1. The absolute difference in TTC between UE and SO, and therefore the benefit of rerouting, is also independent of a_1. However, as equation (2) shows, PoA is dependent on a_1 and, in particular, as $a_1 \to \infty$, $PoA \to 1$. Although PoA therefore captures the benefits of rerouting relative to the total cost of travel, if the free-flow component makes up a significant proportion of that total cost, which in road networks is typically true (Correa et al., 2008: Games and Economic Behavior, 64, 457-469), PoA can mask the still potentially significant absolute benefits of rerouting. We propose an alternative measure to capture these benefits called 'Price of Anarchy Delays' ($PoAD$); equation (3), which compares the relative difference under UE and SO of only those cost components directly affected by the routing of flows. In excluding free-flow costs, $PoAD$ excludes the section of cost that is unavoidable to travellers and which cannot be altered by the routing strategy.

$$PoAD = \frac{TTC^{UE} - \sum_{rs} C_1^{rs}(0) q_{rs}}{TTC^{SO} - \sum_{rs} C_1^{rs}(0) q_{rs}} = \frac{\sum_{rs}\left(\sum_{k=2}^{K} f_k^{UE}\gamma_k + \sum_{k=1}^{K} f_k^{UE}\left[C_k^{UE}\left(f_k^{UE}\right) - C_k^{UE}(0)\right]\right)}{\sum_{rs}\left(\sum_{k=2}^{K} f_k^{SO}\gamma_k + \sum_{k=1}^{K} f_k^{SO}\left[C_k^{SO}\left(f_k^{SO}\right) - C_k^{SO}(0)\right]\right)} \quad (3)$$

Having defined $PoAD$, we now seek to understand whether, like PoA, it is bounded and also how it varies with travel demand. Initially we do this using a single OD network of N parallel links, serving a demand, $q > 0$, with linear link cost functions of the form $C_i(x_i) = a_i + b_i\, x_i$, $\forall i$ where (WLOG) $a_i \le a_{i+1} \forall i$ and all coefficients a_i, b_i are positive. For such networks it transpires that $PoAD$ is subject to the same bounds as those proved by Roughgarden (2003) for PoA. To see this first note from equations (2) and (3) that if $C_1(0) = 0$ then $PoAD = PoA$. Secondly note that, by construction, $PoAD$ is invariant to a transformation of costs $\widetilde{C_i} = C_i + B$ $\forall i$ for any constant cost B. If we transform link costs by setting $B = -a_1$ then link flows and the value of $PoAD$ remain the same, the transformed network produced has $\widetilde{C_1}(0) = 0$ meaning $PoAD = PoA$ and $PoAD$ is therefore subject to the Roughgarden (2003) bounds.

We now examine how and why $PoAD$ varies with travel demand. Under UE, in the defined parallel link networks, at sufficiently low levels of demand, all flow uses only the cheapest link, link 1. This remains true for all values $q > 0$ for which:

$$a_1 + b_1 q \le a_2 \iff q \le \frac{(a_2 - a_1)}{b_1}$$

When demand q exceeds this value link 2 *activates* and both links carry flow at UE. As demand increases subsequent links activate in turn, and this defines a set of demand regimes, until at sufficiently high demand, all network links are active. The situation with SO is similar: increasing demand causes a sequence of link activations though at different demand levels than for UE. For the N parallel link network described, the Mth demand regimes under UE and SO are:

Mth UE Demand Regime: $\sum_{i=1}^{M-1} \frac{a_M - a_i}{b_i} < q \leq \sum_{j=1}^{M} \frac{a_{M+1} - a_j}{b_j}$ Mth SO Demand Regime: $\frac{1}{2}\sum_{i=1}^{M-1} \frac{a_M - a_i}{b_i} < q \leq \frac{1}{2}\sum_{j=1}^{M} \frac{a_{M+1} - a_j}{b_j}$

To examine how $PoAD$ varies with travel demand we consider $N = 2, \dots, 10$ parallel link networks with cost coefficients $a_i = i$ and $b_i = 1$ for $i = 1, \dots, 10$. The variation of $PoAD$ with demand for each network N is shown in Figure 2, with link activation points identified by the vertical lines: green for an activation under UE and red for an activation under SO. Focussing first on the results for $N = 10$, Figure 2 illustrates the same general behaviour as that shown in Figure 1; parity between UE and SO up until the second-shortest path is activated under SO, then a series of peaks and troughs as demand increases and finally a decay back towards parity between UE and SO. The vertical lines show this behaviour arises because links *activate* at different levels of demand under UE and SO, as described above, with the peaks in $PoAD$ occurring at the points of activation of links under UE and the decay towards parity occurring only once all links have been activated. The decay toward parity in the final demand regime is $O(1/q^2)$ for the case of linear link cost functions; a more general result is included in the paper. Figure 2 also illustrates how the variation of $PoAD$ for $N = 10$ subsumes that of its sub-networks; i.e. there is an exact overlap of $PoAD$ for each network $M < N$ up until the activation of the $(M+1)^{th}$ link under SO. Therefore in addition to showing how $PoAD$ varies with travel demand, we have also captured the mechanism, link activations, that drives this variation. We report corresponding results for more general networks (e.g. Sioux Falls) in the paper. Our results also demonstrate that because the set of links in the '*active* network' is dependent on the level of demand, the topology of the supply network effectively changes as demand increases. This is an important consideration when attempting to compare the performance of different supply topologies.

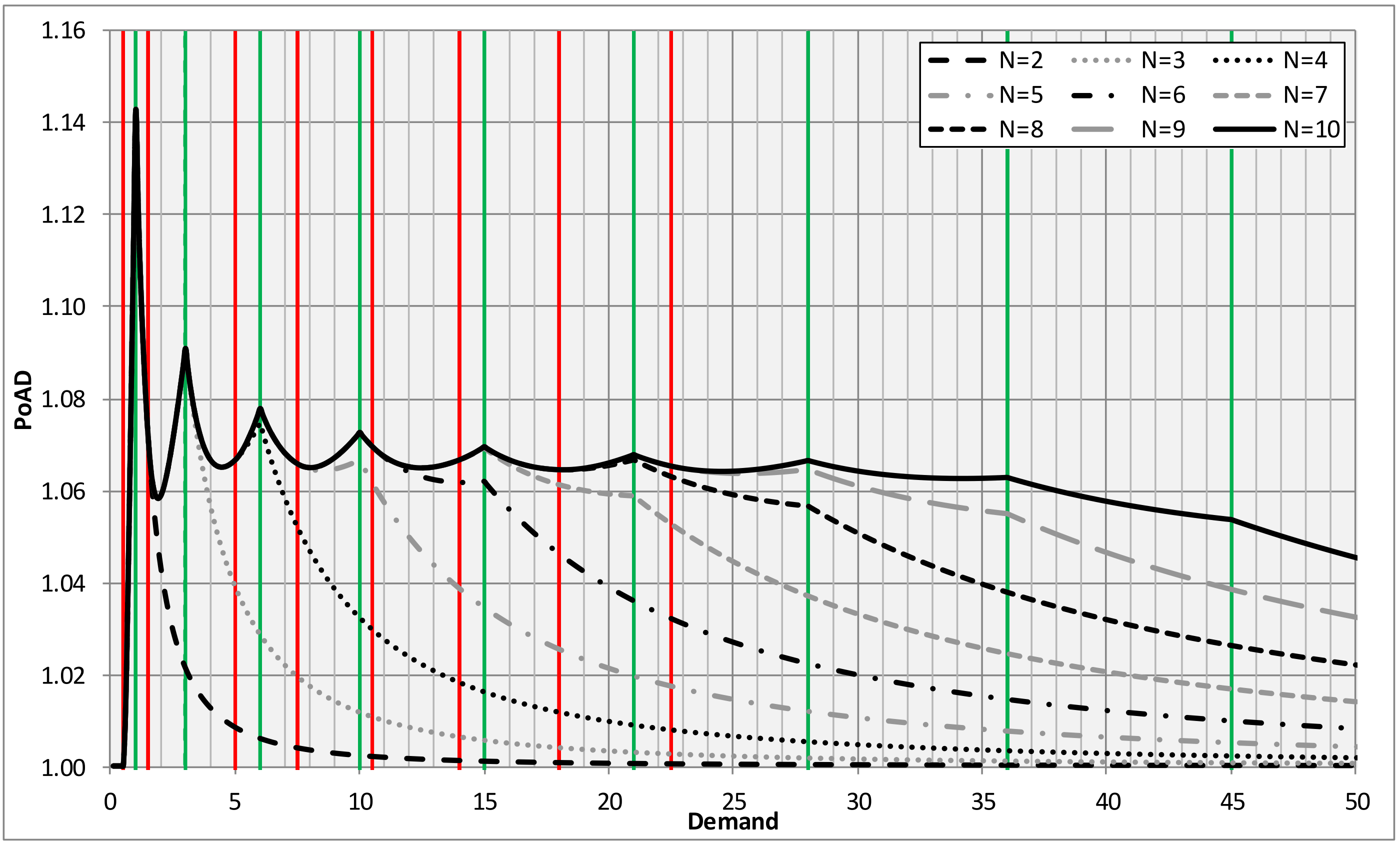


Figure 2: PoAD vs Demand for N-Parallel Link networks and UE (Green) and SO (Red) Link Activation Points

The contribution of this work can therefore be summarised as follows. We have proposed a new measure; Price of Anarchy Delays, for the cost of selfish routing and compared it with the standard Price of Anarchy measure. For parallel link networks with linear cost functions we show $PoAD$ is subject to the same bounds as PoA and we identify the occurrence of link activations and show these underlie the variation of $PoAD$ with travel demand. In the full paper we show these phenomena occur in general networks with multiple ODs and capture the nature of the asymptotic decay of the $PoAD$ back to unity for non-linear cost functions.