

# The Price of Human Nature

LILLIAN TSAI, MIT, USA

VIBHAALAKSHMI SIVARAMAN, MIT, USA

XIAOYUE GONG, MIT, USA

## 1 INTRODUCTION

Finding the best strategy for network and traffic routing has historically been a problem of great importance combining the theoretical aspects of both game theory and computer science. Seminal work by Roughgarden and Tardos framed the routing problem as a non-cooperative game, in which players' selfish routing decisions increase the social welfare cost, i.e., the overall latency of all users in the network. Our paper begins with this selfish model, presenting a brief overview of the original selfish routing problem and model and the limits on the optimality of these algorithms (termed as the "price of anarchy").

Since the advent of the selfish routing model, more complex (and perhaps more realistic) models have emerged, many of which emphasize the need to take into account the complexity of human behaviors. This paper focuses on three of these recent alternative models, namely models that account for altruistic, risk-averse, and diverse behaviors. We present and clarify the findings of these papers in the context of the original selfish routing paper, and demonstrate how these papers' results can be synthesized into a more general framework addressing optimality of routing with various human behaviors or motivations.

## 2 BACKGROUND

This section presents a brief history of traffic routing algorithms, defining the terminology and the traffic routing problem.

### 2.1 The Traffic Routing Problem

Traffic routing problems naturally arise in communication or transportation networks, where users are trying to minimize the latency that they or their data experiences. However, links in the network often becomes *congested* if too many users decide to route their data or cars through that link. Consequently, in these networks, the path each user chooses can affect the travel times of other users. Here, we describe Roughgarden and Tardos' formalization of the problem of minimizing latency as multicommodity flow networks [9, 10], and use Pigou's example network in Figure 1 as a running example. We use this to reason about the deviation from the optimum minimal latency when users make selfish routing decisions.

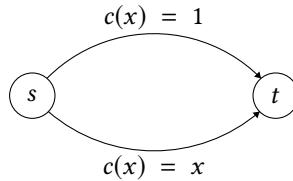


Fig. 1. Pigou's example traffic routing problem, with a demand of  $r_{(s,t)} = 1$

**The input** to a traffic routing problem consists of:

- A network  $G = (V, E)$  of  $|V|$  destinations (e.g., locations or servers) and  $|E|$  links

- A set of  $k$  source-destination pairs  $S = \{(s_1, t_1), \dots, (s_k, t_k)\}$  representing traffic demands
- A rate  $r_i$  of traffic for each  $(s_i, t_i) \in S$  representing the demanded amount of traffic from  $s_i$  to  $t_i$
- A latency cost function  $c$  that assigns a per-edge function  $c_e$  to each edge  $e$  describing how adding traffic (i.e., congestion) to  $e$  affects the time taken to travel across  $e$ . We can also think of  $c$  as assigning per-path costs: for any path  $p$  in the graph

$$c_p(f) = \sum_{e \in p} c_e(f_e)$$

We assume that  $c$  is continuous, nonnegative, and nondecreasing.

In Figure 1, we see a single source-destination input network with an example (linear) cost function with  $r_{(s,t)} = 1$ .

**Solutions** correspond to flow assignments to the set of simple paths  $P_i$  between  $s_i$  and  $t_i$  for all  $i$ . Note that our solutions assume *nonatomic* entities: the flows we find may not be integral. Intuitively, this means that the demand from one  $s_i$  to  $t_i$  is generated by an infinite number of entities in the network, which allows us to reason about continuous, rather than discrete, functions. To describe a flow assignment  $f$ , we can consider  $f_p$ , the flow on a single path  $p \in P_i$  (this adds an equal amount of flow  $f_p$  to all edges in  $p$ ), as well as  $f_e = \sum_p \sum_{e \in p} f_p$ , the flow on edge  $e$  (the sum of flow on all paths that use  $e$ ).

A *feasible* solution given such an input is an assignment of path flows such that the demand from  $s_i$  to  $t_i$  is met:

$$\forall 1 \leq i \leq k, \sum_{p \in P_i} f_p = r_i$$

An *optimal* (feasible) solution given such an input is the feasible flow assignment  $f$  that minimizes the **total weighted cost**  $C(f)$ , where

$$C(f) = \sum_i \sum_{p \in P_i} c_p(f) f_p = \sum_{e \in E} f_e c_e(f_e)$$

Intuitively, we are calculating the cost of each path of a given flow assignment, weighing each path's cost proportional to the amount of flow through the path. More concretely, if we were to let flow represent the routes chosen by (infinitely many) users,  $C(f)$  calculates the average cost over all users. Thus, when minimizing  $C(f)$ , some users may incur more latency so that other users can go faster: the optimal flow is the *socially optimal* solution. Note that there exists an optimal flow  $f^*$  minimizing  $C(f)$  because we assume  $c$  is continuous and the set of feasible flows is compact. We will also refer to the total weighted cost as the **social welfare** cost.

In our running example (Figure 1), a feasible flow is any flow that sends one unit from  $s$  to  $t$  (divided in any fashion between the top and bottom edges). The optimal flow is the flow that sends half the traffic through the lower edge and half through the upper edge: the users on the lower edge only experience a cost of  $1/2$ , while the users on the upper edge experience a cost of  $1$ , making the total weighted cost  $3/4$ .

## 2.2 Coordination Models and the Price of Anarchy

Before we can create algorithms to solve the traffic routing problem, we must first assume a *coordination model* for our traffic network. There are two clear extremes: (1) centralized control, in which some entity (e.g., an air traffic controller) knows all traffic demands and routes accordingly, and (2) decentralization, i.e., a complete *lack* of coordination between entities in the network. In a centralized setting, there is a clear optimal solution, as shown in the previous section. However, in a decentralized and uncoordinated model, the lack of coordination and the exercise of freewill can result in inefficiencies.

**The Price of Anarchy** (PoA) allows us to measure the inefficiencies of a decentralized model given some notion of equilibrium (how the flow assignment is determined in the model), and was first introduced by Koutsoupias and Papadimitriou in 1999 [5]. The PoA is defined as the ratio between the optimal flow and the flow achieved at

equilibrium. (This is similar to how we measured the distance from optimal of an approximation algorithm in a limited computational power model, and of online algorithms in an incomplete information model.)

### 2.3 The Selfish Routing Model

One example of an uncoordinated model is the *selfish routing* model, in which all entities in the network are selfish and choose a route minimizing their individual latency without caring (or knowing) about the effects on other users [10].

The selfish routing model corresponds to flows at a *Wardrop, or Nash equilibrium* [4, 12]. The set of flows in a Nash equilibrium are defined such that for all  $i$  source-destination pairs, all the paths from  $s_i \rightarrow t_i$  have the minimum-possible cost. In other words, the (nonzero) flow paths at Nash equilibrium have equal path costs, and no user decreases her cost by choosing a different path:

$$\forall 0 \leq i \leq k, \forall p_1, p_2 \in P_i \text{ s.t. } f_{p_1} > 0 \text{ and } f_{p_2} > 0, c_{p_1}(f) = c_{p_2}(f)$$

Note that this corresponds exactly to the (polynomial-time) solutions to the following (convex) program:

$$\min_f \left( \sum_e \int_0^{f_e} c_e(t) dt \right) \text{ subject to feasibility constraints}$$

whereas the flow optimizing the social welfare or total cost corresponds exactly to the (polynomial-time) solutions to the following (convex) program:

$$\min_f \left( \sum_e f_e c_e(f_e) \right) \text{ subject to feasibility constraints}$$

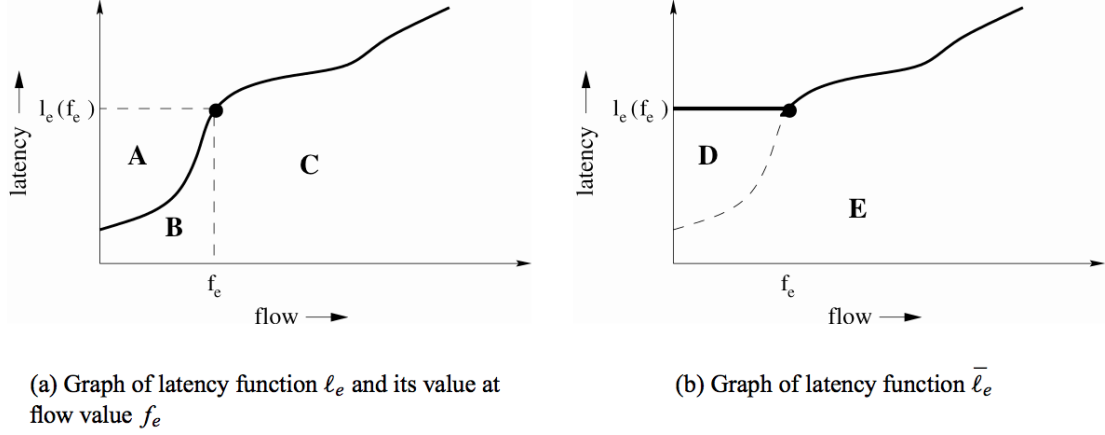
If we revisit our running example in Figure 1, we note that the flow at Nash equilibrium corresponds to a flow that sends the entire unit of traffic through the bottom edge (the 0 flow through the top path has a cost of 1, and the unit flow through the bottom path will have cost 1). Intuitively, each user routing from  $s$  to  $t$  will selfishly choose to take the bottom route because she will reason that the bottom route can have cost no worse than the top route. However, by doing so, the bottom route becomes more congested and leads to a total average cost  $C(f) = 1$ . Thus, in Figure 1, the price of anarchy  $\rho$  is  $\frac{1}{3/4} = \frac{4}{3}$ .

We next describe and present the main results regarding the PoA in this (decentralized) selfish routing model, which will act as a basis to which we will compare traffic routing results in more recently formulated models.

### 2.4 Main Results

**THEOREM 1.** *If  $f$  is a flow at Nash Equilibrium for a given input set  $(G, r, c)$  and  $f^*$  is a feasible flow for  $(G, 2r, c)$ , then  $C(f) \leq C(f^*)$*

**PROOF SKETCH.** In other words, this result suggests that the latency when all users of a network route with only their own interests in mind is utmost the optimal (minimum) latency for the same graph and latency functions with twice as much demand per path.



For any Nash Equilibrium solution  $f$  for  $(G, r, \ell)$  and any feasible solution  $f^*$ , we can draw the figures so that for each edge  $e$  in  $G$ :

The area  $A+B$  represents the cost  $C(f)$  of the Nash Equilibrium solution  $f$  for  $(G, r, \ell)$ . The area  $B+C$  represents the cost  $C(f^*)$  of the feasible solution  $f^*$  for  $(G, 2r, \ell)$ . Since the area  $D+E$  is at most area  $D$  more than area  $B+C$ , and area  $D$  is less than area  $A+B$ , we know that  $D+E-(B+C) \leq D \leq A+B$ . This means that  $B+C \geq D+E-(A+B)$ . On the other hand  $D+E \geq 2(A+B)$ . Therefore, we have that  $C(f^*) = (B+C) \geq 2(A+B) - (A+B) = A+B = C(f)$ . Therefore we have the result that the cost of any Nash equilibrium flow for  $(G, r, \ell)$  is no bigger than represents the cost of any feasible solution for  $(G, 2r, \ell)$ . The main idea of the proof for this theorem lies in that the subset if the area  $C$  is at least as big as the area  $A+B$ . This is always true because the latency  $\bar{\ell}_e$  is a nondecreasing function and thus the graph of latency function gives a shape that looks like a generalized trapezoid (in particular, a right generalized trapezoid lying on the horizontal axis).  $\square$

**THEOREM 2.** *If the edge latency functions are linear in that  $c_e = a_e f_e + b_e$  for every edge  $e \in E$ , then the price of anarchy or  $\rho(G, R, l) \leq 4/3$ .*

**PROOF SKETCH.** The cost of any flow  $f$  under these edge latency functions is  $C(f) = \sum_e a_e f_e^2 + b_e f_e$ .

Let's consider two flows  $f$  and  $f^*$  such that  $f$  is at Nash Equilibrium in  $(G, r, \ell)$  and  $f^*$  is globally optimal for the same. We first consider optimally routing the first  $r/2$  demand across all source-destination pairs. It turns out that  $f/2$  is optimal for  $(G, r/2, \ell)$  when the edge latency functions are linear. This can be derived from the fact that paths with non-zero flow at a Nash Equilibrium have the same path latency while paths with non-zero flows at the global optimum have the same marginal cost of increasing the flow. Now, if we look at the cost  $C(f/2)$  of routing this in terms of the cost of routing the flow  $f$  at Nash Equilibrium, we notice that  $C(f/2) = \sum_e \frac{1}{4} a_e f_e^2 + \frac{1}{2} b_e f_e \geq \frac{1}{4} C(f)$  from the above cost expression. Thus, in other words, routing the first  $r/2$  optimally has a latency that is at least one-fourth of the latency of the Nash Equilibrium flow.

This leaves the remaining  $r/2$  that needs to be routed optimally to route  $f^*$  fully. To reason about this, let's look at a small  $\delta r_i$  increase in flow from  $s_i$  to  $t_i$  that already carries  $x$  units of flow. For a convex cost function, we expect the increase in latency to be at least  $\delta r_i l'(x)$  where  $l'$  is the minimum marginal increase in  $C$ . If we consider starting at the optimal flow  $f/2$  for the  $r/2$  demand and increasing the flow on each path by a small

$\delta r_i$ , the subsequent increase in latency across all paths can be summed as  $\sum_{i=1}^k l'(f/2) \delta r_i$ . But, for linear edge latency functions, the marginal increase in latency on every edge at  $f/2$  is exactly the latency of that edge at  $f$ . Thus,  $l'_e(f/2) = l_e(f)$ . Thus, setting  $\delta = 1$  and increasing the rate by  $r_i/2$  on every  $s_i$ , the overall increase in latency is at least  $\frac{1}{2} \sum_{i=1}^k l(f) r_i = \frac{1}{2} C(f)$ . The last part is by definition of  $C(f)$ .

In essence, routing the first  $r/2$  demand optimally costs at least  $C(f)/4$  and the next  $r/2$  when augmented, costs at least another  $C(f)/2$ . In total, the optimal costs at least  $\frac{3}{4} C(f)$ . In other words, the flow at Nash Equilibrium has cost utmost  $\frac{4}{3} C(f^*)$  where  $f^*$  is the flow achieving optimal latency.  $\square$

### 3 ALTERNATIVE MODELS

In this section, we present and compare a subset of recent coordination models for the traffic routing problem against the selfish routing model. These models propose a more nuanced (and perhaps more accurate) description of human behavior. [XXX TODO more summary once we have more insights?](#)

#### 3.1 Altruism and Spite

The first alternative model we consider is that proposed by Chen and Kempe in 2008 [2], which assumes that users are “not entirely selfish.” Chen and Kempe note that social experiments from both economic and psychology have shown humans do not behave rationally in a selfish manner; instead, our behavior is better modeled as either altruistic or malicious (spiteful). Their model proposes a simple way to capture how people make decisions based upon how much cost (latency) a particular decision will cost other users; if someone is spiteful, she will want to increase their cost, and if she are altruistic, she will want to decrease their cost.

**3.1.1 Formalization.** The formal Chen and Kempe model introduces a per-user *altruism* coefficient  $\beta$  and a new cost function  $c_p^\beta$  for all paths  $p$ :

$$c_p^\beta(f) = \sum_{e \in p} c_e(f_e) + \beta \sum_{e \in p} f_e c'_e(f_e)$$

where  $c_e(\cdot)$  is the cost function from the selfish routing setting, and  $c'_e(\cdot)$  is the derivative with respect to  $f_e$ .

Note that the first term is exactly the cost used in the selfish routing model (and thus is equivalent to an altruism coefficient of  $\beta = 0$ ). The second term corresponds to the derivative of the social welfare cost on  $p$  and is weighed by  $\beta$ ; we use the derivative, rather than the value, of the social welfare cost on  $p$  because each user only controls an infinitesimally small amount of the flow. Thus, if we were to use the value, a single user’s choice would have no effect on the social welfare cost! Instead, a user can account for how she will affect the social welfare cost via the rate at which her choice of path affects other users.

If  $\beta$  is negative, a user is spiteful: we know that adding a little more flow to  $p$  will increase the social welfare cost of taking  $p$  (the derivative  $c'_e$  is positive), and since we negate this value, this lowers the user’s perceived cost of taking  $p$ . Conversely, if  $\beta$  is positive, a user is altruistic: increasing flow increases the social welfare cost on  $p$  and also the user’s perceived cost of taking  $p$ . We assume that  $\beta$  ranges from -1 (extremely spiteful) to 1 (extremely altruistic), where  $\beta = 0$  corresponds to selfishness.

All analysis of the model assumes a particular distribution  $\psi$  of  $\beta$  for all users. Furthermore, we can compare this model to the selfish model using the PoA as a measure of inefficiency because the altruistic model still achieves Nash equilibrium for any  $\psi$  and cost function  $c_p^\beta$ : given any  $\psi$ ,  $c_p^\beta$  are continuous in the choice of path  $p$ , and we can apply the result of Theorem 1 of Mas-Colell [?]. Nash equilibrium is achieved at the flow solutions to the

program

$$\min_f \sum_e \int_0^{f_e} c_e^\beta(t) dt \text{ subject to feasibility constraints}$$

We next present Chen and Kempe's core results about the PoA of arbitrary networks when  $\psi$  is uniform (all users have the same  $\beta$  value), and briefly mention their results of non-uniform  $\psi$  in parallel-link networks.

**3.1.2 Uniformly Distributed Altruism.** We first consider the case where  $\psi$  is uniformly distributed, such that  $\beta$  and therefore  $c_p^\beta$  is the same for each user. We additionally assume that users tend to be altruistic, i.e.,  $\beta > 0$ .

**THEOREM 3.** *For any input network  $G$ , demand rates  $r$ , and a uniform distribution  $\psi$  with  $\beta \in (0, 1]$ , if  $c_e$  is nondecreasing and convex for all  $e$ , then the price of anarchy is bounded by*

$$\rho(G, r, c, \psi) \leq \frac{1}{\beta}$$

**PROOF SKETCH.** Consider the two (convex) objectives  $NE(f)$  and  $C(f)$  where  $NE$  is the objective of a flow at Nash equilibrium, and  $C$  represents the social welfare cost:

$$\begin{aligned} NE(f) &= \sum_e \int_0^{f_e} c_e^\beta(t) dt = \sum_e \int_0^{f_e} c_e(t) + \beta t c'_e(t) dt \text{ (by definition of } c_e^\beta) \\ C(f) &= \sum_e f_e c_e(f_e) = \sum_e \int_0^{f_e} (t c_e(t))' dt = \sum_e \int_0^{f_e} c_e(t) + t c'_e(t) dt \end{aligned}$$

It is clear that for any feasible flow  $f$ ,  $NE(f) \leq C(f) \leq \frac{NE(f)}{\beta}$  because  $\beta \in (0, 1]$ . We now let  $\hat{f}$  be the flow at Nash Equilibrium and  $f^*$  be the flow at optimum social welfare. Because these are the optimal flows for their respective objectives, we know that  $C(\hat{f}) \leq \frac{NE(\hat{f})}{\beta} \leq \frac{NE(f^*)}{\beta} \leq \frac{C(f^*)}{\beta}$ , proving that  $\rho(G, R, c, \psi) \leq \frac{1}{\beta}$   $\square$

**3.1.3 Uniformly Distributed Spite.** Chen and Kempe then address the problem of spite: how (uniformly) spiteful can users be before the PoA becomes infinite?

XXX TODO something about anarchy value and given class of cost functions

**3.1.4 Arbitrarily Distributed Altruism.** Chen and Kempe go on to extend their analysis to when users have an arbitrary distribution  $\psi$  of altruism (with no spiteful users) in *parallel link networks*. We briefly mention their results here, but direct the reader to the paper for a more detailed explanation. Their main result mirrors that of the PoA with uniform altruism:

**THEOREM 4.** *Given any parallel link network  $G$ , demand rates  $r$ , altruism density function  $\psi$  with average altruism  $\bar{\beta}$  and non-negative support, and convex and non-decreasing cost functions  $c_e$ ,*

$$\rho(G, r, c, \psi) \leq \frac{1}{\bar{\beta}}$$

One interesting observation is that this theorem leads to the following corollary: because the theorem applies for a distribution in which a rate of  $\bar{\beta}$  of users are completely altruistic and  $1-\bar{\beta}$  users are completely selfish, this is equivalent to  $\bar{\beta}$  of the population being under centralized (coordinated) control:

**COROLLARY 5.** *Given any parallel link network  $G$ , demand rates  $r$ , altruism density function  $\psi$  with average altruism  $\bar{\beta}$  and non-negative support, and convex and non-decreasing cost functions  $c_e$ , if  $\bar{\beta}$  fraction of traffic is controlled by a central authority, then*

$$\rho(G, r, c, \psi) \leq \frac{1}{\bar{\beta}}$$

This is exactly the result proven by Roughgarden in 2004 [8], and provides an interesting connection between models that include altruism and models of centralized/coordinated traffic control.

XXX Maybe address spite (not possible to include negative support)

### 3.2 Risk Aversion

The second model we consider accounts for the tendency of users to pick routes with less variation in latency even if it comes at the cost of some added latency on the paths chosen. This increase in latency can be quantified as the *price of risk-aversion* which is the worst-case ratio of the latency or cost at a risk-averse Nash Equilibrium to that at a risk-neutral Nash equilibrium or one where users are indifferent to variations in the latency itself.

**3.2.1 Formalization.** The **formal model** introduced in Lianes et.al [6] defines a risk aversion coefficient  $\gamma$  that quantifies the users' tendency to prefer paths with less variability. A higher  $\gamma$  means that one is more risk averse. The edge costs  $c_e(x_e)$  now have a deterministic part or  $l_e(x_e)$  and a noise modelled by a random variable  $\xi_e(x_e)$ . The latter is assumed to be independent across edges and has expectation 0 and variance  $v_e(x_e)$  for  $x_e$  flow allowing us to sum them up over a path. To simplify the analysis, the model also defines  $\kappa$  to bound the variance-to-mean latency ratio. In other words,  $v_e(x_e) \leq \kappa l_e(x_e)$ .

The cost function or the **objective** that every user is now optimizing for is of the form

$$c_p(f) = \sum_{e \in p} l_e(f_e) + \gamma \sum_{e \in p} v_e(f_e)$$

This function is assumed to be non-decreasing. Intuitively, the mean is identical to the original price of anarchy formulation while the second term accounts for variance. Minimizing this implies that we want to minimize the variance depending on the value of the risk coefficient itself.

If the maximum cost  $C(x)$  across some set of nash equilibria flows  $x$  for the given problem instance that is restricted to some family of inputs and a fixed  $\kappa$ , the **price of risk-aversion** is now defined as the ratio  $C(x)/C(z)$ . Here  $C(z)$  is the cost associated with some risk-neutral nash equilibrium  $z$ . In the following section, we look at bounding this price of risk-aversion. Note that this can be viewed as contributing a multiplicative factor to the price of anarchy in the overall change to the latency or cost of the system. We first prove a more basic result from an older paper on this topic [7] and then proceed to the main result on the price of risk-aversion for a special family of latency functions that are  $(\lambda, \mu)$  smooth. Note that additional results can be proved for special classes of graph topologies which can be found in the original paper [6].

#### 3.2.2 Main Results.

**THEOREM 6.** *If a flow  $f$  is at a risk averse nash equilibrium and  $f^*$  is any other flow, then  $fC(f) \leq f^*C(f)$ .*

**PROOF.** By definition, any flow at Nash Equilibrium routes on paths with minimum cost or only sends flow on a given path if its cost is less than the cost of sending the same flow on some other path. Thus, for a fixed demand and fixed path costs determined by  $C(f)$ ,  $f^*$  differs from  $f$  in atleast moving some  $\epsilon$  flow from one lower cost path to another higher cost one which increases the total overall cost. Consequently, let's say we have a flow  $x$  that is a risk averse Nash Equilibrium for the cost function  $c_e = l_e(x_e) + \gamma v_e(x_e)$ . If  $z$  is a risk neutral Nash Equilibrium, it is still feasible for the risk averse mean-variance cost function, but is not the equilibrium in that scenario. By the above description and costs written as the sum across edges, we have

$$\sum_{e \in E} x_e(l_e(x_e) + \gamma v_e(x_e)) \leq \sum_{e \in E} z_e(l_e(x_e) + \gamma v_e(x_e))$$

□

**Definition 7.** A function  $l : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is  $(\lambda, \mu)$ -smooth around  $x \in \mathbb{R}_{\geq 0}$  if  $yl(x) \leq \lambda yl(y)_{\mu} xl(x)$  for all  $y \in \mathbb{R}_{x \geq 0}$

Intuitively, what this means is XXX add some intuition

**THEOREM 8.** *The set of instances with latency functions  $l_{e \in E}$  that are  $(1, \mu)$ -smooth around any risk-averse nash equilibrium  $x_e$  for all  $e \in E$  have price of risk-aversion  $\leq \frac{(1 + \gamma\kappa)}{(1 - \mu)}$*

**PROOF SKETCH.** seaprate into edges with flow less than  $z_e$   
edges with flow greater than  $x_e$  - inequality to be subtracted □

### 3.3 Diverse in Interests

The third class of alternative models we consider is a generalization of the altruistic model and the risk-averse model that we have discussed in previous sections. This class is characterized by the diverse selfish behavior by its heterogeneous agents. Each agent pursues their own different selfish goal, resulting in a routing solution of the whole network.

Diverse selfish routing models are useful because they help us understand how we can leverage policies and natural diversity of goals in a network to increase the social welfare and efficiency of the network as a whole. For example, tolls can help increase the social welfare. For example, Beckmann et. al. showed that tolls can help implement the social optimum as an equilibrium, when agents all have the same goal towards a linear combination of time and money [1].

However, there is some ambiguity in measuring the optimality of any outcome of the whole network with diverse selfish behavior, because by definition, the objective function has changed from the objective of selfish routing with no agent heterogeneity and thus only one criterion. There are thus multiple reasonable ways to characterize the social welfare of a diverse routing network. We will discuss the model adopted by Cole, Lianas and Nikolova and their newly published results in 2018 [3].

**3.3.1 Model.** We have the same routing network with multiple source-destination pairs and continue with all our previous notations, except that we have included two criteria the players consider in the objective function.

Each agent wants to minimize their own cost, which is a sum of two terms associated with two criteria. Let  $\ell_P$  denote the cost of the first criterion (e.g., the latency) over some path  $P = (s_i, t_i)$ , and  $\sigma_P$  be the cost of the second criterion, referred to as the *deviation function*. Then given a routing  $f$  of the network, the cost of that path is given by  $\ell_P + \gamma \cdot \sigma_P = \sum_{e \in P} \ell_e(f_e) + \sum_{e \in P} \sigma_e(f_e)$ , where  $\gamma$  is the *diversity parameter*.

Note that the latency function has all the properties as we assumed in previous sections, while the deviation function  $\sigma_e(x)$  is assumed to be continuous by not necessarily non-decreasing. However, the function  $\ell_e + \gamma \cdot \sigma_e$  must be non-decreasing. These assumptions are consistent with our previous risk averse model in Section 3.2, because if  $\sigma_e$  models the variance, then  $\sigma_e$  could be decreasing in the flow.

Cole et. al. measures the effect of diversity against the resulting flow of a homogeneous agent population of the same size. The homogeneous agent population has the single diversity parameter  $\bar{r} = \int r f(r) dr$ .

For a discrete distribution of  $n$  discrete values  $r_1^k, \dots, r_n^k$ , the demand  $d_k$  is a vector  $d_k = (d_1^k, \dots, d_n^k)$  where each  $d_i^k$  denotes the total demand of commodity  $k$  with diversity parameter  $r_i^k$  and  $d^k$  denote commodity  $k$ 's total demand  $d^k = \sum_{i=1}^n d_i^k$ . For a heterogeneous equilibrium flow vector  $g$ , the *heterogeneous total cost* of commodity  $k$  is denoted by  $C^{k,ht}(g) = \sum_{j=1, \dots, n} d_j^k c^{k,r_j^k}(g)$ , where  $c^{k,r_j^k}(g)$  denotes the common cost at equilibrium  $g$  for players of diversity parameter  $r_j^k$  in commodity  $k$ . The heterogeneous total cost of  $g$  is then  $C^{ht}(g) = \sum_{k \in K} C^{k,ht}(g)$ .

For the corresponding homogeneous equilibrium flow  $f$ , i.e. the instance with diversity parameter  $\bar{r}^k$ , where  $\bar{r}^k$  denotes the average diversity parameter for commodity  $k$ , players of commodity  $k$  share the same cost  $c^{\bar{r}^k}(f)$ . Then, the homogeneous total cost of commodity  $k$  under  $f$  is  $C^{k,hm}(f) = d_k c^{\bar{r}^k}(f)$ , and the homogeneous total cost of  $f$  is  $C^{hm}(f) = \sum_{k \in K} C^{k,hm}(f)$ .



**3.3.2 Results.** Let  $g$  denote an equilibrium flow for the heterogeneous agent population and  $f$  an equilibrium flow for the corresponding homogeneous agent population. Let  $C^{ht}(g)$  denote the cost of flow  $g$  and  $C^{hm}(f)$  the cost of flow  $f$ .

A *multi-commodity network* is consistent with all our previous models. We also introduce the definition of a *single-commodity network* as a network whose edges all belong to some single source-destination path as only these edges are going to be used by the equilibria and thus all other edges can be discarded. We present the following main results.

**Definition 9.** A directed  $s - t$  network  $G$  is *series-parallel* if it consists of a single edge  $(s, t)$ , or it is formed by the series or parallel composition of two series-parallel networks with terminals  $(s_1, t_1)$  and  $(s_2, t_2)$ , respectively.

The theorem below essentially states that for single-commodity networks, diversity is always helpful in a single-commodity series parallel network.

**THEOREM 10.** For any  $s - t$  series-parallel network  $G$  with a single commodity, we have  $C^{ht}(g) \leq C^{hm}(f)$ .

**PROOF SKETCH.** The key observation is that since  $f$  and  $g$  route the same amount of flow from the unique source to the unique sink, there must be a path  $P$  where  $f$  sends no less flow along than  $g$  does. Since the network is series-parallel, for every edge  $e$  in  $P$ ,  $f_e \geq g_e$ . This is true because a path in a series-parallel network can be broken up into some series-parallel parts in series and some series-parallel parts in parallel, which recursively breaks down to a simple series of edge(s). Hence for any  $r \in [0, r_{\max}]$ , we have  $c_p^r(f) \geq c_p^r(g)$ . This simultaneously also means that  $g$  could route more flow on  $P$  but it doesn't, implying that there is no incentive for  $g$  to switch the flow it sends on other paths to path  $P$  under the same diversity parameter. Therefore  $c^r(g) \leq \sum_{e \in P} \ell_e(g_e) + r \sum_{e \in P} \sigma_e(g_e)$ . Then  $C^{ht}(g) \leq \sum_{i=1}^k d_i(\sum_{e \in P} \ell_e(g_e) + r_i \sum_{e \in P} \sigma_e(g_e)) = \ell_p(g) + \bar{r} \sigma_P(g)$  which is exactly the cost of the homogeneous equilibrium flow  $f$ .  $\square$

**THEOREM 11.** For any  $s - t$  non-series-parallel network  $G$  with a single commodity, there exists cost functions  $C$  for which  $C^{ht}(g) > C^{hm}(f)$ .

**PROOF SKETCH.** If  $G$  is not series-parallel then the Braess graph can be embedded in it [11], and there are edge functions such that heterogeneous equilibrium flow has a larger cost than homogeneous equilibrium flow. Detailed example can be found in [3].  $\square$

This theorem essentially states that for single-commodity networks, diversity is always helpful only in a series-parallel network. Together with Theorem 10, we know that the series-parallel structure is a sufficient and necessary condition for diversity to always be helpful.

Now we discuss our main results for multiple commodity network. We use *average-respecting demand* to refer to the property that for any commodity  $i, j : \bar{r}^i = \bar{r}^j$ .

A multi-commodity network  $G$  can be decomposed in subnetworks  $G_i$ 's that each contains all the vertices and edges of  $G$  that belong to a simple  $s_i - t_i$  path for commodity  $i$ . WLOG, we assume these  $G_i$  are acyclic.

**Definition 12.** A multi-commodity network  $G$  is *block-matching* if for every  $i$ ,  $G_i$  is series-parallel, and for every  $i, j$ ,  $G_i$  and  $G_j$  are block-matching, respectively.

The next theorem states that for multi-commodity networks, diversity is always helpful on any block-matching network with average-respecting demand.

**THEOREM 13.** For any  $k$ -commodity block-matching network with average-respecting demand,  $C^{ht}(g) \leq C^{hm}(f)$ .

**PROOF SKETCH.** The key idea is that for any block representation of a subnetwork  $G_i = s_i B_1 v_1 \dots v_{b_i-1} B_{b_i} t_i$  for some commodity  $i$ , because  $G$  is block-matching, for any block  $B_j$  connecting  $v_{j-1}$  and  $v_j$ , any other commodity

$j$  either contains block  $B_j$  as a block in its block representation or contains none of the edges of  $B_j$ . This implies that under any routing of the demand, either all of  $j$ 's demand goes through  $B_j$  or none of it does. So the total traffic routed by  $f$  and  $g$  are the same from  $v_{j-1}$  to  $v_j$ . So if restricted to the block, the cost of the heterogeneous equilibrium is less than or equal to that of the homogeneous equilibrium; then the theorem is a result of summing over every block of all commodities.  $\square$

**THEOREM 14.** *For any  $k$ -commodity network, if diversity helps for every instance on  $G$  with average-respecting demand, we have  $C^{ht}(g) \leq C^{hm}(f)$ , then  $G$  is a block-matching network.*

**PROOF SKETCH.** The proof is by contradiction. First, by our Theorem 11 for single-commodity network, each subnetwork  $G_i$  in our multi-commodity network must be a series-parallel network, otherwise we can use the same counterexample as for Theorem 11. Then since we can prove that any two commodities  $i$  and  $j$ , any block of  $G_i$  and any block of  $G_j$  either have the same terminals and direction or their terminals has no intersection, we know that  $G$  is block-matching. The detailed proof assumes the conditions does not hold and constructs demand and edge functions where diversity hurts to contradict the assumption. See [3].  $\square$

This theorem essentially states that for multi-commodity networks, diversity is always helpful only on any block-matching network with average-respecting demand. Together with Theorem 13, we know that the block-matching structure is a sufficient and necessary condition for diversity to always be helpful in a multi-commodity network.

## 4 DISCUSSION

### 4.1 Importance of Human Understanding

### 4.2 Future Work

Clearly the models we present are a miniscule subset of all potential models for human behavior; furthermore, they are coarse-grained and oversimplistic in comparison with the complexity of the human brain. As understanding of the neurophysiological aspects of human behavior improves, we hope to see a matching evolution in the precision and accuracy of these models for traffic routing as well.

## REFERENCES

- [1] BECKMANN, M., MCGUIRE, C., AND WINSTEN, C. *Studies in the economics of transportation*. Research memorandum. Published for the Cowles Commission for Research in Economics by Yale University Press, 1956.
- [2] CHEN, P.-A., AND KEMPE, D. Altruism, selfishness, and spite in traffic routing. In *Proceedings of the 9th ACM Conference on Electronic Commerce* (New York, NY, USA, 2008), EC '08, ACM, pp. 140–149.
- [3] COLE, R., LIANEAS, T., AND NIKOLOVA, E. When does diversity of agent preferences improve outcomes in selfish routing? In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18* (7 2018), International Joint Conferences on Artificial Intelligence Organization, pp. 173–179.
- [4] HAURIE, A., AND MARCOTTE, P. On the relationship between nash and cournot and wardrop equilibria. *Networks* 15, 3, 295–308.
- [5] KOUTSOPIAS E., P. C. Worst-case equilibria. *STACS 1999. Lecture Notes in Computer Science* (1999).
- [6] LIANEAS, T., NIKOLOVA, E., AND STIER-MOSES, N. E. Risk-averse selfish routing. *Mathematics of Operations Research* (2018).
- [7] NIKOLOVA, E., AND STIER-MOSES, N. E. A mean-risk model for the traffic assignment problem with stochastic travel times. *Operations Research* 62, 2 (2014), 366–382.
- [8] ROUGHGARDEN, T. Stackelberg scheduling strategies. *SIAM Journal on Computing* 33, 2 (2004), 332–350.
- [9] ROUGHGARDEN, T. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005.
- [10] ROUGHGARDEN, T., AND TARDOS, E. How bad is selfish routing? *J. ACM* 49, 2 (Mar. 2002), 236–259.
- [11] VALDES, J., TARJAN, R. E., AND LAWLER, E. L. The recognition of series parallel digraphs. In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 1979), STOC '79, ACM, pp. 1–12.
- [12] WARDROP, J. Some theoretical aspects of road traffic research. *Inst Civil Engineers Proc* (1952).