Routing Network Flow Among Selfish Agents

Aaron Johnson

May 14, 2004

Abstract

We consider the problems posed by routing flow in a network populated by self-interested agents. Standard node-cost and edge-cost network models are compared and a mapping between them is described. The existence of Nash equilibria when flow cannot be split is established in several cases. Braess's paradox is shown to exist when the number of users is finite and flow is unsplittable. Finally, optimizing the flow routing with polynomial and capacity cost functions is shown to be hard to compute and to approximate.

1 Introduction

In this paper, we consider problems arising from the situation in which a group of agents sends traffic within a network. Generally, a cost function on the amount of flow is associated either with each vertex or with each edge, and each agent wishes to minimize its own cost, while the network administrator wishes to minimize the total cost. We examine these models and show they are equivalent for the types of problems we consider. We extend results for the existence and uniqueness of Nash equilibria when agents cannot split their flow. In this *Atomic Noncooperative Network* scenario, we examine the well-known Braess's paradox and provide examples of its existence. Finally, we examine the problem of computing the optimal flow routing with unsplittable flow.

The game theoretical aspects of network flow routing have been of growing interest to the computer science community ([20] , [11] , [7] , [12], [6]). This is a part of a larger focus on noncooperative behavior in computational contexts that has become more relevant with the growth of the Internet [17]. Other research directions in this area include network flow control [10], multicast transmissions [5], and peer-to-peer systems [2].

In selfish routing, several models have been proposed and results achieved. Roughgarden thoroughly explores the "price of anarchy" in his thesis [20]. This is the worst-case cost ratio between the Nash equilibrium and optimal routing in a directed network with an infinite number of agents, each with a negligible amount of flow. He also examines the problem of designing a network topology for such agents and achieves hardness results. Mitigating this effect in an

efficient way with taxes is explored in [3]. The existence and uniqueness of equilibria in networks with a finite number of users, each of whom can split their flow among multiple paths, is examined and established under certain conditions in [16]. Allocating capacity in a network when users behave selfishly is considered in [13] and an efficient algorithm to achieve the optimal equilibrium.is presented. Nisan and Ronen introduced the problem of helping agents route their traffic efficiently when the vertex or edge transit cost is constant but known only to the owner in [15]. An efficient centralized solution to this problem is given in [9], and an efficient distributed algorithm for this problem in the context of BGP routing is given in [6].

2 NetworkModel

In the literature, the network is represented either as a graph with cost functions on the edges or as a graph with cost functions on the vertices. We choo'se to model the network with costs on the edges, since this is more common. It will be shown here that there exists a map M going from each graph G in the vertex-cost model to a graph G' in the edge-cost model, such that for every flow in G there exists a flow in G' with the same cost. Similarly, there exists such a map from each graph in the edge-cost model to a graph in the vertex-cost model.

Let us consider a model in which one agent controls a certain amount of flow in a network. The network is represented by a directed graph $G = (V, E)$. The agent is represented by the sequence (s, t, r), where $s, t \in V$ are the source and destination vertices, respectively, and $r \in \mathbb{R}$ is the amount of flow that agent wishes to send. This model and the theorems about it easily generalize to the situation in which there are multiple agents.

For the vertex-cost model, let $c : V \to \{f : \mathbb{R} \to \mathbb{R}\}$ be a function from the vertices of the graph to a function on the real numbers representing the cost function of a vertex. Then a model instance in vertex-cost is represented by the sequence (G, c, a). For clarity, we will denote the cost function on a vertex $v \in V(c(v))$ as c_v.

For the edge-cost model, let $c : E \to \{f : \mathbb{R} \to \mathbb{R}\}$ be a function from the edges of the graph to a function on the real numbers representing the cost function of a vertex. A model instance in edge-cost is represented by the sequence (G, c, a). Again, for clarity we will refer to the cost function on an edge $e \in E(c(e))$ as c_e.

Definition 2.1 (Neighbor set). Given a graph $G = (V, E)$, the *incoming neighbor set* of a vertex $v \in V$ is defined as the set $\Gamma^-(v) = \{u \mid (u, v) \in E\}$. Similarly, the *outgoing neighbor set* of v is defined as the set $\Gamma^+(v) = \{u \mid (v, u) \in E\}$.

Definition 2.2 (Flow). In general, a *flow* is a function $f : E \to \mathbb{R}$ such that

$$\forall_{e \in E} f(e) \geq 0$$

$$\forall_{v \in V - \{s,t\}} \sum_{u \in \Gamma^-(v)} f((u,v)) = \sum_{u \in \Gamma^+(v)} f((v,u))$$

We will further restrict the definition of a flow with the condition that it respect the amount of flow sent by an agent:

$$\sum_{v \in \Gamma^+(s)} f((s,v)) = \sum_{v \in \Gamma^-(t)} f((v,t)) = r$$

Definition 2.3 (Node cost). Given (G, c, a), an instance of the vertex-cost model, and g, a flow in that instance. Then $f : V \to \mathbb{R}$, the flow through a vertex, is known as the *vertex flow* and is defined as

$$f(v) = \begin{cases} \sum_{w \in \Gamma^-(v)} g((w,v)) = \sum_{w \in \Gamma^+(v)} g((v,w)) & \text{if } v \in V - \{s,t\} \\ 0 & \text{otherwise} \end{cases}$$

This definition is well-defined because the first and second sums are equal in a flow for all vertices that are not the source and destination.

Then let $c(g)$ be the cost of the flow, defined as

$$c(g) = \sum_{v \in V} f(v) c_v(f(v))$$

Definition 2.4 (Edge cost). Given (G, c, a), an instance of the edge-cost model, and f, a flow in that instance. Then $c(f)$, the cost of the flow, is defined as

$$\sum_{e \in E} f(e) c_e(f(e))$$

Theorem 2.5. *Given an instance $I = (G, c, a)$ of the node-cost model, there exists an instance $I' = (G', c', a')$ of the vertex-cost model for which given any flow f in I there exists a flow f' in I' such that the costs of the two flows are equal.*

Proof. The instance I' is constructed as follows.

Let $V' = \{v^-, v^+ \mid v \in V\}$. Let $E' = \{(u^+, v^-) \mid (u,v) \in E\} \cup \{(v^-, v^+) \mid v \in V\}$. Let $G' = (V', E')$.

Let

$$c'((u,w)) = \begin{cases} c(v) & \text{if for some } v \in V,\ u = v^- \wedge w = v^+ \\ 0 & \text{otherwise} \end{cases}$$

Let $a' = (s^+, t^-, r)$.

Finally, let

$$f'((u,v)) = \begin{cases} f(w,x) & \text{if } u = w^+ \wedge v = x- \\ \sum_{w \in \Gamma^-(u)} f(w,u) = & \\ \sum_{w \in \Gamma^+(v)} f(v,w) & \text{if } u = x^- \wedge v = x^+ \text{ for some } x \in V \end{cases}$$

It is easy to check that this is a flow in I'. Then

$$\begin{aligned} c(f) &= \sum_{v \in V} f(v) c_v(f(v)) \\ &= \sum_{v \in V} f'((v^-, v^+)) c_v(f'((v^-, v^+))) \\ &= \sum_{v \in V} f'((v^-, v^+)) c_{(v^-, v^+)}(f'((v^-, v^+))) \\ &= \sum_{e \in E} f'(e) c_e(f(e)) \\ &= c(f') \end{aligned}$$

□

Theorem 2.6. *Given an instance $I = (G, c, a)$ of the edge-cost model, there exists an instance $I' = (G', c', a')$ of the vertex-cost model for which given any flow f in I there exists a flow f' in I' such that the costs of the two flows are equal.*

Proof. The instance I' is constructed as follows.

Let $V' = V \cup \{e \mid e \in E\}$. Let $E' = \{(u, e), (e, v) \mid e = (u, v) \in E\}$. Let $G' = (V', E')$.

Let

$$c'(v) = \begin{cases} c(e) & \text{if } v = e \in E \\ 0 & \text{otherwise} \end{cases}$$

Let a'=(s, t, r).

Let

$$f'((u, v)) = \begin{cases} f(v) & \text{if } v = e \text{ for some } e \in E \\ f(u) & \text{if } u = e \text{ for some } e \in E \end{cases}$$

It is easy to check that this is a flow. Then

$$\begin{aligned} c(f) &= \sum_{e \in E} f(e) c_e(f(e)) \\ &= \sum_{e \in E} f'(e) c_e(f'(e)) \\ &= \sum_{v \in V} f(v) c_v(f(v)) \\ &= c(f') \end{aligned}$$

□

3 Nash Equilibria

In selfish routing, as in other noncooperative games, finding flows for which the system is in equilibrium is useful in understanding the system. We expect that

the system will stay at such flows if they are ever reached, so their performance could largely determine the system's performance. We would generally like to know when equilibrium flows exist, if they are unique, and how their cost compares to the minimum flow cost. Results in these areas exist for three classes of the edge-cost network model.

The first class is the classical model, in which there are infinitely many agents controlling an infinitesimally small amount of flow that they can send along only one path. This model was first studied in the 50s ([1], [21]) in the context of traffic research. It was more recently the primary model for network traffic studied by Roughgarden [20]. It is assumed that the cost functions on the edges are nonnegative, continuous, and nondecreasing. In this case, it is a result of [1], and later [4], that a Nash equilibrium flow exists, and is unique in the sense thats its total cost is equal to the total cost of any other Nash flow.

In the second class there are finitely many agents, each of whom can split the flow he sends to his destination along different paths. This type of game was considered by Orda, Rom and Sihmkin in [16] in slightly more generality, allowing the cost functions on each edge to be different for each user. With the assumption that the edge cost functions of each user are nonnegative, continuous, differentiable and convex, they sketched a proof that this game is a concave n-person game of the type described in [18]. As a result every such game has a Nash equilibrium.

The uniqueness results from Orda, et al., for this type of game require more assumptions. However, uniqueness is taken in the stronger sense that every Nash flow is equal. The situations in which a Nash flow is unique are:

- If the agents are *symmetric* then the NE is unique. Agents are called symmetric if the share the same edge cost functions, have identical source-destination pairs, and control the same amount of flow.
- If cost functions of the agents are *diagonally strict convex* (DSC), then the game possesses a unique Nash equilibrium. The DSC condition is defined in [18], and sufficient conditions for the cost functions are given by Orda et al. that are simpler to check.
- If the cost functions of the user satisfy certain reasonable conditions and the network consists of a set of parallel links, then the NE is unique. The conditions are:
  - Each edge function J_l^i for user i on edge l is a function of the total flow on the edge and the component of that flow contributed by user i.
  - J_l^i is increasing in both arguments.
  - $\frac{\partial J_l^i}{\partial f_l^i}$ is increasing in f_l^i and f_l.

In the third class there are finitely many agents, and each agent must send all of his flow along one path from the source to the destination. Libman and Orda

term these games *Atomic Noncooperative Networks* (ANN) [12]. In this game the agent's strategies are the paths from their sources to their destinations, a finite set in a finite network. This restriction to a finite set of combinations of strategies, or strategy profiles, seems to make the analysis harder, and the results for this scenario are less broad.

Libman and Orda, in [12], prove the existence of Nash equilibrium flows in parallel-link ANNs with some conditions on the edge cost functions. The conditions are:

- The cost functions are the same for every edge.
- The cost functions are increasing with the flow
- The cost function for every edge l is only defined over the interval $[0, C_l]$. C_l is a real value that represents the capacity of the edge.

They also describe a simple algorithm to determine if the Nash flow for the network is unique.

In 1973, Rosenthal [19] introduced a class of games termed *congestion games*, for which he proved a Nash equilibrium always exists. Later it was shown [22] that these games are isomorphic to the more well-known class of *potential games*. The definition of congestion games and proof of the existence of Nash equilibria will be given here since it is then used achieve two results about ANNs.

Definition 3.1 (Congestion games). Given a set of agents $N = \{1, \ldots, n\}$ and a set of primary factors $T = \{1, \ldots, t\}$. The strategy set Σ_i of each agent i is a subset of the power set, $\Sigma_i \subseteq 2^T$. Thus each strategy consists of a set of primary factors. A cost function $c_k : \mathbb{R} \to \mathbb{R}$ is associated with each $k \in T$, and takes the number of agents using k. Given a strategy profile σ, the cost $\pi_i(\sigma)$ to each agent i is the sum of the costs of each of the factor in the strategy for that agent. That is, let $\sigma \in \Sigma_1 \times \cdots \times \Sigma_n$ be a strategy profile. Let $x_k(\sigma) = |\{\sigma_i \mid k \in \sigma_i, 1 \leq i \leq n\}|$ be the number of agents using k in σ. Then

$$\pi_i(\sigma) = \sum_{k \in \sigma_i} c_k(x_k(\sigma))$$

Theorem 3.2 (Rosenthal). *All congestion games possess at least one pure-strategy Nash equilibrium.*

Proof. Let x_j^i be 1 if the i^{th} agent plays the j^{th} strategy $P_j \in 2^T$, or zero if it does not. A solution to the following problem must exist:

$$\begin{cases} \text{Minimize} & \sum_{k=1}^{t} \sum_{y=0}^{x_k} c_k(y) & & \\ \text{Constrained by} & \sum_{j=1}^{2^{|T|}} x_j^i = 1 & i = 1, \ldots, n & \\ & x_k - \sum_{i}^{n} \sum_{j \mid k \in P_j} x_j^i = 0 & k = 1, \ldots, t & \\ & x_j^i = 0 \quad \text{or} \quad 1 & i = 1, \ldots, n & j = 1, \ldots, 2^{|T|} \end{cases}$$

Suppose a solution $\{x_j^i, x_k\}$ to this system is not a Nash equilibrium. Then there exists some agent a playing strategy P_j such that for some other strategy

P_l

$$\sum_{\substack{k \in P_l \\ k \notin P_j}} c_k(x_k + 1) < \sum_{\substack{k \notin P_l \\ k \in P_j}} c_k(x_k)$$

Let $\{x_j^{'i}, x_k^{'}\}$ be the variable set induced by a switching strategies from P_j to P_l. Then

$$\begin{aligned}\sum_{k=1}^{t} \sum_{y=0}^{x_k^{'}} c_k(y) &= \sum_{k=1}^{t} \sum_{y=0}^{x_k} c_k(y) + \sum_{\substack{k \in P_l \\ k \notin P_j}} c_k(x_k + 1) - \sum_{\substack{k \notin P_l \\ k \in P_j}} c_k(x_k) \\ &< \sum_{k=1}^{t} \sum_{y=0}^{x_k} c_k(y)\end{aligned}$$

This is a contradiction. Therefore, there must exist a Nash equilibrium and thet solution to the above program is one. □

Congestion games are similar to ANNs. In fact, if all of the agents in an ANN control the same amount of flow, a mapping exists to the set of congestion games.

Theorem 3.3. *Every ANN in which the agents have equal flow is a congestion game.*

Proof. Given an ANN instance, $A = (G, c, a)$, in which each agent controls r units of flow, construct a congestion game $C = (N, T, \Sigma, c)$. Let the agents in C be the agents in A, $N = a$. Let the edges of G be the primary factors of C, $T = E$. Let the strategies of each agent be every simple path from the source to the destination, $\Sigma_i = \{s_i - t_i paths\}$. Let the cost function c_k of C take the values at every nonnegative integer multiple of r from 0 to n. That is, $c_k(i) = c_k(ir), 0 \leq i \leq n$.

These games are equivalent. They have the same number of players. For each player an isomorphism between the strategies in C and those in A exists, as long as we assume that players in A will never consider paths that are not simple, since they can only increase their cost. The payoffs that map to each other under these isomorphisms are the same, as well, because every feasible flow in A can only put some nonnegative integer multiple of r on any given edge, since each agent controls r flow and can only choose one simple path for that flow. □

Corollary 3.4. *Every ANN in which every agent controls the same amount of flow has a Nash equilibrium.*

Proof. This follows directly from 3.3 and 3.2. □

This arguments can be extended to a special case of when the agents have different amounts of flow.

Lemma 3.5. *Given the ANN $I = (G, c, a)$ in which the functions in the range of c are identical and linear, and the paths in G between every source-destination pair of a are of the same length m. Let I' be the ANN identical to I except for the agent flows, which have been scaled by some nonnegative factor ρ. That is, $\forall_i r_i^{'} = \rho r_i$. If some flow f is a Nash equilibrium flow in I, then the flow $f'(e) = \rho f(e)$ is also a Nash equilibrium .*

Proof. Let the identical linear cost function of each edge be $ax + b$. Let r_i be the flow controlled by agent i. Let f be a Nash equilibrium flow, and f_i the component of the flow contributed by agent i. Let P_i be the set of paths in G between the source and destination of i, with p_i^* the one chosen in f. Then by definition of NE the cost to i of $c_i(f)$ must be less than that of choosing any other path, or

$$\forall_{p_i \in P_i} \sum_{e \in p_i^*} r_i(af(e) + b) \leq \sum_{e \in p_i} r_i(af(e) + b)$$

Since the paths are of equal length m, this implies that

$$\forall_{p_i \in P_i} \sum_{e \in p_i^*} f(e) \leq \sum_{e \in p_i} f(e)$$

Then for the flow $f'(e) = \rho f(e)$

$$\begin{aligned}
\forall_{p_i \in P_I} c_i(f') &= \sum_{e \in p_i^*} \rho r_i(a\rho f(e) + b) \\
&= \rho r_i \left[a\rho \sum_{e \in p_i^*} f(e) + bm \right] \\
&\leq \rho r_i \left[a\rho \sum_{e \in p_i} f(e) + bm \right] \\
&= \sum_{e \in p_i} \rho r_i(a\rho f(e) + b)
\end{aligned}$$

Therefore p_i^* is the preferred path for i in f'. □

Theorem 3.6. *If every agent controls a rational amount of flow, the edge cost functions are identical and linear, and all of the simple paths between a given source-destination pair are the same length, a Nash equilibrium must exist.*

Proof. Given an ANN $I = (G, c, a)$, scale the flow controlled by every agent so that the amounts are all integral. By Lemma 3.5, a Nash equilibrium flow found in the game will be an NE flow in the original game.

Let r_i be the amount of flow controlled by every agent i. Let x_j^i be r_i if the i^{th} agent plays the j^{th} strategy $P_j \in 2^T$, or zero if it does not. A solution to

the following problem must exist:

$$\begin{cases} \text{Minimize} & \sum_{k=1}^{t}\sum_{y=0}^{x_k} c_k(y) & & \\ \text{Constrained by} & \sum_{j=1}^{2^{|T|}} x_j^i = r_i & i=1,\ldots,n & \\ & x_k - \sum_i^n \sum_{j|k\in P_j} x_j^i = 0 & k=1,\ldots,t & \\ & x_j^i = 0 \quad \text{or} \quad r_i & i=1,\ldots,n & j=1,\ldots,2^{|T|} \end{cases}$$

Let $M = \{k \mid k \in P_l \wedge k \notin P_j\}$ and $N = \{k \mid k \notin P_l \wedge k \in P_j\}$. Suppose a solution $\{x_j^i, x_k\}$ to this system is not a Nash equilibrium. Then there exists some agent a playing strategy P_j such that for some other strategy P_l

$$\sum_{k\in M} c_k(x_k + r_i) < \sum_{k\in N} c_k(x_k) \tag{1}$$

Let $\{x_j^{'i}, x_k^{'}\}$ be the variable set induced by a switching strategies from P_j to P_l. Then

$$\begin{aligned} \sum_{k=1}^{t}\sum_{y=0}^{x_k^{'}} c_k(y) &= \sum_{k=1}^{t}\sum_{y=0}^{x_k} c_k(y) + \sum_{k\in M}\sum_{y=0}^{r_i-1} c_k(x_k + r_i - y) - \sum_{k\in N}\sum_{y=0}^{r-1} c_k(x_k - y) \\ &= \sum_{k=1}^{t}\sum_{y=0}^{x_k} c_k(y) \quad + \\ &\quad \sum_{y=0}^{r_i-1}\left[\sum_{k\in M}(a(x_k+r_i)+b) - \sum_{k\in N}(ax_k+b) - ay|M| + ay|N|\right] \end{aligned}$$

Because all paths are the same length, $|M| = |N|$, and (1) implies that

$$\sum_{k=1}^{t}\sum_{y=0}^{x_k^{'}} c_k(y) < \sum_{k=1}^{t}\sum_{y=0}^{x_k} c_k(y)$$

This is a contradiction. Therefore, there must exist a Nash equilibrium and at least one is a solution to the above program. □

4 Braess's Paradox

An interesting phenomenon in the classical model of network routing is Braess's paradox [14], in which adding an edge to a network can actually increase the cost of an equilibium flow. The paradox shows that improving the performance of a given network with noncooperative agents is not just a simple matter of adding edges. We will show that this paradox can also occur with the number of agents is finite, and each agents must choose one path for his flow. Note that in the classical model, an equilibrium flow must exist, and the cost of any two equilibrium flows is the same, so the concept of this paradox is well-defined. In

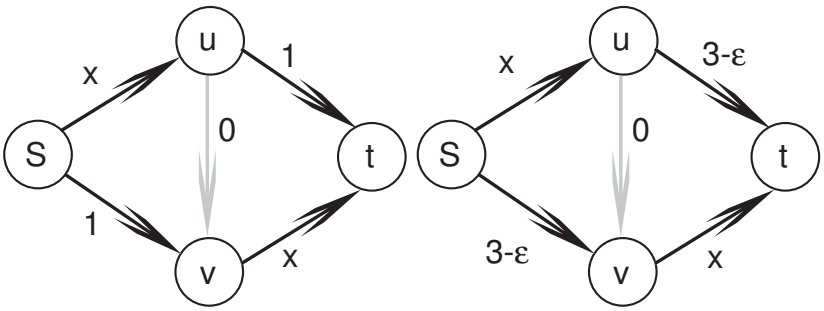

Figure 1: Instance of Braess' paradox with infinite and finite agents

the model with finite users and unsplittable flow, however, an equilibrium may not exist, and even if it does it may not have a unique cost.

To give an example of Braess's paradox in the classical model, we will first state without proof a simple characterization of a flow at equilibrium due to Wardrop ([20]).

Theorem 4.1 (Wardrop's Principle). *A flow f feasible for instance (G,c,a) is at Nash equilibrium if and only if for every* $i \in \{1, \ldots, k\}$ *and* $P_1, P_2 \in P_i$ *with* $f_{P_1} > 0, c_{P_1}(f) \leq c_{P_2}(f)$.

A graph which can suffer from Braess's paradox is shown on the left in Figure 1. The cost functions in the total flow x are indicated next to the edges. Adding the light edge (u, v) can increase the cost of the equilibrium flow. Let there be one unit of flow in the graph from s to t. If (u, v) is not present, the equilibrium and optimal flows coincide, placing $\frac{1}{2}$ unit of flow on the path (s, u, t) and $\frac{1}{2}$ unit on the path (s, v, t). The total cost of this flow is $\frac{3}{2}$. After the edge (u, v) is added, the flow which puts two units of flow on the path (s, u, v, t) and no flow on any other path has a cost of two on every $s - t$ path, so by Theorem 4.1 is an equilibrium flow. The cost of this flow is $1(1) + 0 + 1(1) = 2 > 3/2$, so adding (u, v) increased the equilibrium cost.

A graph on which Braess's paradox can occur with finite users is shown on the right in Figure 1. The cost functions are indicated as before. Let there be two agents in this network, each sending one unit of flow from s to t. If the edge (u, v) is not present, then the unique equilibrium flow is again equal to the optimal flow, and one agent puts his flow on the path (s, u, t) while the other puts his flow on the path (s, v, t). The cost of this flow is $2[1(3-\epsilon)+1(1)] = 8-2\epsilon$. If the edge (u, v) is present, the unique equilibrium flow is when both agents place their flow on the path (s, u, v, t). There are not simple conditions to verify this is the unique equilibrium, but there are only four cases to enumerate. The cost of this flow is $2[1(2) + 1(2)] = 8 > 8 - 2\epsilon$, so adding (u, v) again increased the equilibrium cost.

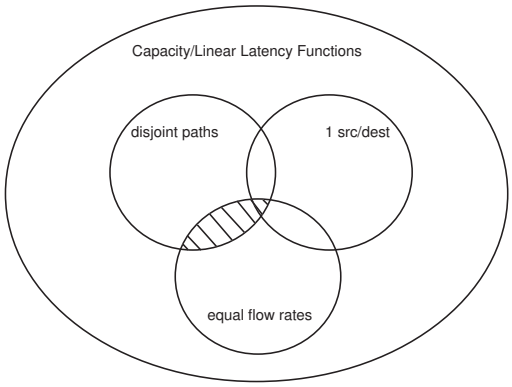

Figure 2: Routing problem and subproblems

5 Complexity of Unsplittable Flow Routing

Since ANN games are a relatively recent model in which to study network performance, a natural question to ask is the complexity of finding the optimal flow in such networks. That is, how hard it is to determine the flow that minimizes the sum of the individual costs of each agent, in the case that each agent must choose just one path for his flow. If the agents could freely communicate and traded off money and network performance equally, you could even expect this flow to be achieved through trading in a noncooperative scenario.

The complexity of this problem has a great deal to do with the type of cost functions allowed at the edges. In the case that the cost of a link is constant, the problem becomes no harder than the all-pairs shortest path problem. Since a number of equilibrium results apply only to networks with convex cost functions, and these reflect expected real-world network performance, we will consider the complexity of certain convex cost functions. In particular, we will look at linear cost functions, which can be no harder than polynomials of higher degree, and capacity cost functions, of the form $1/(mu - x)$, which can describe the delay of $M/M/1$ queues with capacity μ [20].

5.1 Complexity of Optimal Flow

Optimal unsplittable flow routing is in general an NP-hard problem. It seems that this complexity stems from both the number of paths to consider between every source-destination pair and the different flow amounts the every agent may control. In light of this, we have divided the problem space as shown in Figure 2. The universe is the set of all problems in which the edge cost functions are all either linear, and of the form $c_e(x) = ax + b$, or capacity functions of the form $c_e(x) = 1/(\mu - x)$. The circle labeled "disjoint paths" indicates the set of problem instances in which all paths between a source-destination pair are edge-disjoint. The circle labeled "1 src/dest" indicates the set of problem

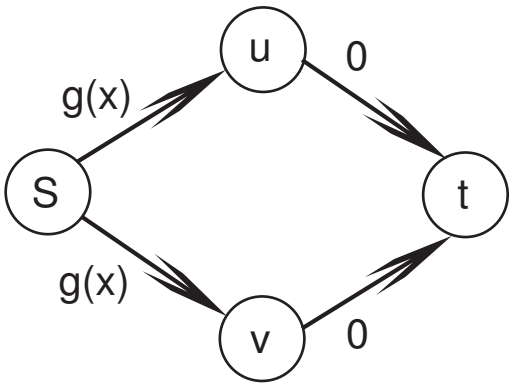

Figure 3: PARTITION graph

instances in which every agent shares the same source-destination pair. The circle labeled "equal flow rates" indicates the set of all problem instances in which the amount of flow controlled by every agent is the same.

We will show that the unshaded subproblems of Figure 2 are NP-hard by reduction from PARTITION and DISJOINT CONNECTING PATHS [8]. The PARTITION reduction will be on problems in the intersection of the "disjoint" and "1 src-dst" problems. The DISJOINT reduction will be on problems in the intersection of the "equal flow" and "1 src-dst" problems. Together these imply the NP-hardness of the entire unshaded region. The complexity of the shaded problem space is unknown.

Theorem 5.1. *Let S be the set of all instances of the edge-cost model in which one source-destination pair is shared by all agents, all the paths between that pair are edge-disjoint, and the edge cost functions are either all linear or all capacity functions. The problem of determining the minimum cost of a flow in an instance of S is NP-hard.*

Proof. Given an instance of PARTITION, a set of n positive integers $B = \{a_1, \ldots a_n\}$. Let $\sum_{i=1}^{n} a_i = A$.

Construct an instance of S with linear cost functions as follows:

- Let $G = (V, E)$ be the graph $V = \{s, t, u, v\}$, $E = \{(s, u), (u, t), (s, v), (v, t)\}$. Figure 3 shows this graph.
- Let the cost functions on $\{(u, t), (v, t)\}$ be 0. Let the functions on $\{(s, u), (s, v)\}$ be $g(x) = x$.
- Let there be n agents who share the source-destination pair (s, t). Let the amount of flow controlled by agent n_i be a_i.

  If under a flow f, m units of flow are routed on the path (s, u, t), the cost of f is $c(f) = m^2 + (A - m)^2 = 2m^2 - 2Am + a^2$. This is minimized when $m = A/2$. Therefore B has a subset of sum $A/2$ iff the optimal routing has cost $A^2/2$. This reduction can easily be done in time $O(n)$.

Construct an instance of S with capacity functions in a similar way, only changing the function $g(x)$. Let $\delta = \min(\min_{1\leq k\leq n}(a_k), \min_{1\leq i,j\leq n}(|a_i - a_j|))$. let the functions on $\{(s,u),(s,v)\}$ be

$$g(x) = \begin{cases} 1/(\frac{A}{2} + \epsilon - x), & 0 < \epsilon < \delta \quad \text{if} \quad x < \frac{A}{2} + \epsilon \\ \infty & \text{otherwise} \end{cases}$$

Then there is a finite cost routing iff some subset sum of B is $A/2$.

The value δ can be calculated in $O(nlogn)$ time by sorting the values of B and finding the minimum of the first value and the minimum of the successive differences. □

Theorem 5.2. *Let S be the set of all instances of the edge-cost model in which all of the agents have equal flow and the edge cost functions are either all linear or capacity functions. The problem of determining the minimum cost of a flow in an instance of S is NP-hard.*

Proof. Given a graph $G = (V, E)$ and a set of disjoint vertex pairs $\{(s_1, t_1), \ldots (s_n, t_n)\}$ that is an instance of DISJOINT CONNECTING PATHS.

We will construct a problem instance (G',c,a) in the vertex-cost model, since DISJOINT is stated in terms of vertices. We then rely on Theorem 2.5 to ensure that there is in instance in S with flows of the same cost.

Construct this vertex-cost instance with linear cost functions as follows:

- Let $G' = G$.
- Let the cost function $g(x)$ of every vertex $v \in V$ be $g(x) = x - 1$.
- Let there be n agents with source-destination pairs $\{(s_1, t_1), \ldots, (s_n, t_n)\}$. Each agents controls one unit of flow.

The cost at every vertex is zero at $x = \{0, 1\}$ and positive for $x > 1$. Since each agent controls one unit of flow, there is a flow with cost zero iff there is a set of vertex-disjoint paths between all n source-destination pairs.

Construct another instance with capacity functions in the same way, except for every $v \in V$, let the cost function be

$$g(x) = \begin{cases} \frac{1}{1+\epsilon-x}, & 0 < \epsilon < 1 \quad \text{if} \quad x < 1 + \epsilon \\ \infty & \text{otherwise} \end{cases}$$

This instance has a finite flow iff there is a set of vertex-disjoint path between every source-destination pair.

Both constructions can be done easily in polynomial time. □

Corollary 5.3. *Let S be the set of all instances of the edge-cost model in which all of the agents share the same source-destination pair, have equal flow and the edge cost functions are either all linear or all capacity functions. The problem of determining the minimum cost of a flow in some instance of S is NP-hard.*

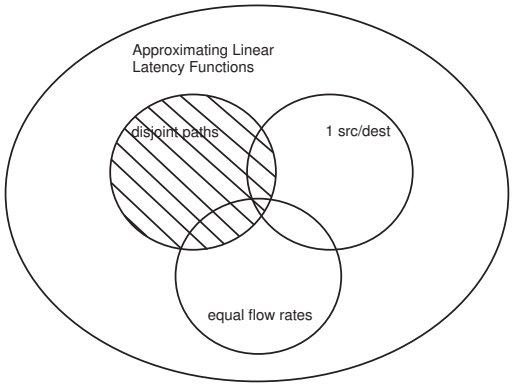

Figure 4: Approximating linear routing

Proof. Make the same constructions as in 5.2. Then add a vertex s and edges connecting that vertex to every previous source vertex, $\{(s, s_i) \mid 1 \leq i \leq n\}$. Similarly, add vertex t and edges connecting to it from every previous destination, $\{(t_i, t) \mid 1 \leq i \leq n\}$. Let (s, t) be the source-destination pair for each agent. In this construction there are just as many edges leaving s and entering t as there are agents. Also, the cost functions on the set $\{s_i, t_i\}_i$ now count in the cost since they are no longer sources or destinations. Since every agent is identical, there is a zero (finite) flow in the instance with linear (capacity) cost functions iff there is a set of disjoint paths connecting each s_i and t_i. This additional work can be done in $O(n)$. □

5.2 Complexity of Approximating Optimal Flow

Since optimal unsplittable flows are hard to compute, we would like to investinage methods to efficiently compute approximations. Unfortunately, some of the reductions given to show their hardness also show that often no constant-factor approximation is feasible. For capacity cost functions, the results carry over directly. Figure 5.2 shows the hardness of approximation with linear costs, where the unshaded problem space in NP-hard and the complexity of the shaded problem space is unknown.

Corollary 5.4. *Let S be the set of all instances of the edge-cost model in which all of the agents share the same source-destination pair, have equal flow and the edge cost functions are either all linear or all capacity functions. The problem of approximating the minimum cost of a flow by a constant factor ρ in an instance of S is NP-hard.*

Proof. Given an instance of DISJOINT CONNECTING PATHS, make the construction of an instance $I \in S$ as described in Corollary 5.3. As described in that proof, with linear cost functions the answer to DISJOINT is YES iff the minimum cost flow of I is zero. With capacity cost functions the answer to

DISJOINT is YES iff the minimum cost flow of I is finite. If some algorithm approximates the minimum cost flow by a factor ρ, with linear functions it must return zero, and with capacity functions it must return a finite value in I, iff the DISJOINT answer is YES. □

Corollary 5.5. *Let S be the set of all instances of the edge-cost model in which one source-destination pair is shared by all agents, all the paths between that pair are edge-disjoint, and the edge cost functions are all capacity functions. The problem of approximating the minimum cost of a flow in an instance of S by a constant factor ρ is NP-hard.*

Proof. Given an instance of PARTITION, create the construction of an instance $I \in S$ as described in Theorem 5.1. As stated in that theorem, the PARTITION problem's answer is YES iff there is a finite flow in I. Therefore, an algorithm which approximates the minimum cost flows in S by ρ must return a finite value iff the PARTITION answer is YES. □

6 Conclusions and Future Work

In general, the Atomic Noncooperative Network model performs worse and is more complex than the classical model or the model with finite users and splittable flow. Equilibria may not exist, and may not be unique when they do. ANNs still suffer from Braess's paradox. They are also very hard to optimize, and even approximations to the optimal appear to be very difficult.

The work done so far has only begun to answer question posed by this model. More general existence and uniqueness NE results could be obtained, as well as examples of ANNs without equilibria. This could lead to a comparison, in the line of inquiry of [20], between the performance of Nash and optimal flows in ANNs.

The classical and ANN models could be viewed as a regions in a range over the number of selfish users with unsplittable flow. It seems that as the number of users decreases, the equilibrium performance also decreases. It would be very interesting to investigate this relationship, as well as the same relationship when the users can split their flow. Also, a model in which small groups of users cooperate to improve the total group welfare could be very appropriate in some contexts, and the performance of such a system might be explored.

The complexity results achieved for optimizing ANN flows leave some large gaps in the problem space. Filling those gaps would help us understand the source of complexity in this problem. Also, algorithms for the approximation of the optimal flow, especially in the linear case, look feasible

References

[1] M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation.*, Yale University Press, 1956.

[2] Chiranjeeb Buragohain, Divyakant Agrawal, Subhash Suri. "A Game Theoretic Framework for Incentives in P2P Systems." IEEE P2P 2003, Linkoping, Sweden. Sept. 1-3, 2003.

[3] Richard Cole, Yevgeniy Dodis, Tim Roughgarden. "Pricing Networks with Selfish Routing". Presented at the *Workshop on Economics of Peer-to-peer Systems*, July 28, 2003.

[4] S. C. Dafermos and F. T. Sparrow. "The traffic assignment problem for a general network." *Journal of Research of the National Bureau of Standards, Series B*, 73B(2):91-118, 1969.

[5] Joan Feigenbaum, Christos Papadimitriou, and Scott Shenker. "Sharing the Cost of Multicast Transmissions." in *Proceedings of the 2000 ACM Symposium on Theory of Computing*, 2000.

[6] Joan Feigenbaum, Christos Papadimitiou, Rahul Sami. "A BGP-Based Mechanism for Lowest-Cost Routing." *Proceedings of the 2002 ACM Symposium on Principles of Distributed Computing*, 2002.

[7] Eric Friedman, "Genericity and Congestion Control in Selfish Routing" . Working paper, (2003).

[8] Michael R. Garey and David S. Johnson. "Computers and Intractibility: A Guide to the Theory of NP-Completeness." W.H. Freeman and Company, New York. 1979.

[9] John Hershberger and Subhash Suri. "What is an edge worth?." 42nd Annual Symposium on Foundations of Computer Science, 2001.

[10] Yannis A. Korilis and Aurel A. Lazar. "On the Existence of Equilibria in Noncooperative Optimal Flow Control." Journal of the ACM, May 1995.

[11] Elias Koutsoupias and Christos Papadimitriou. "Worst-case Equilibria". In *Proceeding of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, G. Meinel and S. Tisod eds., pp. 387-396,, Vol. 1563, Lecture Notes in Computer Science, Springer-Verlag, Trier, Germany, March 1999.

[12] Lavy Libman and Ariel Orda. "Atomic Resource Sharing in Noncooperative Networks." In *Proceedings of the INFOCOM '97*, 1997.

[13] Lavy Libman and Ariel Orda. "The Designer's Perspective to Atomic Noncooperative Networks." *IEEE/ACM Transactions on Networking, Vol. 7 Issue 6*, December 1999.

[14] J.D. Murchland. "Braess's paradox of traffic flow." *Transportation Research*, 4:391-394, 1970.

[15] Noam Nisan and Amir Ronen. "Algorithmic Mechanism Design." *Games and Economic Behavior (35)*, 2001. Extended abstract appeared at the *Thirty-First Annual ACM Symposium on Theory of Computing*, 1999.

[16] Ariel Orda, Raphael Rom, and Nahum Shimkin. "Competitive Routing in Multi-User Communication Networks." *IEEE/ACM Transactions on Networking*, 1993.

[17] Christos Papadimitriou. "Algorithms, Games and the Internet." In *Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing*, pages 749-753, 2001.

[18] J. Rosen. "Existence and uniqueness of equilibrium points for concave n-person games." *Econometrica*, vol. 33, pp. 520-534, July 1965.

[19] Robert W. Rosenthal. "A Class of Games Possessing Pure-Strategy Nash Equilibria." *International Journal of Game Theory*, 2, pp. 65-67, 1973.

[20] Tim Roughgarden, "Selfish Routing." PhD thesis, 2002.

[21] J. G.Wardrop. "Some theoretical aspects of road traffic research." In *Proceedings of the Institute of Civil Engineers, Pt. II*, volume 1, pages 325-378, 1952.

[22] Mark Voorneveld, Peter Borm, Freek van Megen, Stef Tijs. "Congestion Games and Potentials Reconsidered." *International Game Theory Review*, Vol. 1, Nos. 3-4, pp. 283-299, 1999.