Selfish Routing

Simon Fischer

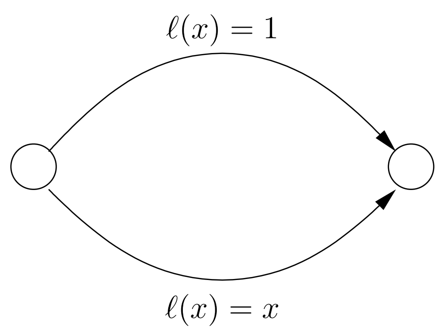
December 17, 2007

1 Selfish Routing in the Wardrop Model

This section is basically a summery of [7] and [3].

1.1 Some Examples

1.1.1 Pigou's Example

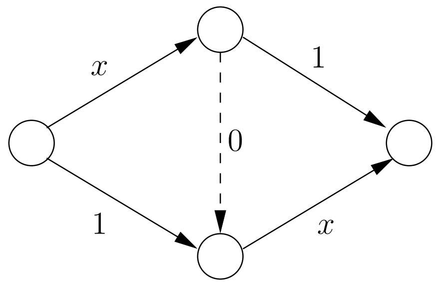


- Optimal solution: split flow $\frac{1}{2}/\frac{1}{2}$ via both edes. Then,

$$OPT = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

- But then: Agents on top edge are unhappy and start to migrate to bottom edge.
- Equilibrium (all agents happy): Total flow on bottom edge. Then $cost = 1$

1.1.2 Braess' Paradox



- First ignore dashed edge.
- Optimal solution: split flow $\frac{1}{2}$ via both paths. Then,

$$OPT = \frac{1}{2} \cdot \left(\frac{1}{2} + 1\right) + \frac{1}{2} \cdot \left(1 + \frac{1}{2}\right) = \frac{3}{2}$$

- Now, insert dashed edge:
- Optimal solution stays the same
- Equilibrium: All agents use zigzag path. Then, $cost = 1 \cdot (1 + 1) = 2$.

1.2 The Wardrop Model

We are given a

- directed graph $G = (V, E)$
- k *commodities* with source-sink pairs s_i, t_i and flow demands r_i, $i \in [k]$, normalise $r = \sum_{i \in [k]} r_i = 1$.
- denote set of paths between s_i and t_i by $\mathcal{P}_i$, let $\mathcal{P} = \cup_{i \in [k]} \mathcal{P}_i$ (for simplicity, let the $\mathcal{P}_i$ be disjoint)
- *latency functions* on the edges $\ell_e : [0, 1] \mapsto \mathbb{R}_{\geq 0}$ (non-negative, non-decreasing, differentiable)
- The triple (G, r, ℓ) is an *instance* of the routing problem.

Agents induce flow and latency:

- *flow vector* $(f_P)_{P \in \mathcal{P}}$, $f_e = \sum_{P \ni e} f_P$.
- a flow is *feasible* if it satisfies the flow demands and flow conservation
- edge latency: $\ell_e(f) = \ell_e(f_e)$
- path latency: $\ell_P(f) = \sum_{e \in P} \ell_e(f)$
- *edge cost*: $c_e(f) = f_e \cdot \ell_e(f)$.

- total *social cost* (average latency):

$$C(f) = \sum_{e\in E} c_e(f) = \sum_{e\in E} f_e \ell_e(f) = \sum_{P\in\mathcal{P}} f_P \ell_P(f)$$

So far, the above is basically a flow model. There exist various game theoretic variants based on this model. In the Wardrop model, we usually assume that flow is controlled by an infinite numer of agents, each responsible for an infinitesimal fraction. Agents strive to minimise their own latency. A flow is at equilibrium if no agent has an incentive to shift their flow unilaterally:

Definition 1 (Nash equilibrium)**.** *A feasible flow f is at a Nash equilibrium iff for every commodity $i \in [k]$, every two paths $P_1, P_2 \in \mathcal{P}_i$ with $f_{P_1} > 0$, and every $\delta \in [0, f_{P_1}]$ the following holds. For the modified flow $\tilde{f}$ where a flow amount of δ is shifted from P_1 to P_2, i. e.,*

$$\tilde{f}_P = \begin{cases} f_P - \delta & \text{if } P = P_1 \\ f_P + \delta & \text{if } P = P_2 \\ f_P & \text{otherwise,} \end{cases}$$

we have $\ell_{P_1}(f) \leq \ell_{P_2}(\tilde{f})$.

Using continuity and monotonicity of the latency functions and letting δ tend to 0 we obtain from this definition a very useful characterisation [5], commonly called *Wardrop's principle* or *Wardrop equilibrium.*

Lemma 1 (Wardrop's principle)**.** *A feasible flow vector f is at a Nash equilibrium iff for every commodity $i \in [k]$, every two paths $P_1, P_2 \in \mathcal{P}_i$ with $f_{P_1} > 0$, we have $\ell_{P_1}(f) \leq \ell_{P_2}(f)$.*

Furthermore, if f is a Nash equilibrium we see that all used paths of the same commodity have minimal latency whereas unused paths may have larger latency. Let $L_i(f)$ denote this minimal latency for commodity i and a Nash equilibrium f. Then,

$$C(f) = \sum_{i\in[k]} r_i \cdot L_i(f) \ .$$

To measure the degradatin of performance due to the lack of central coordination, we use the so-called *coordination ratio* or *Price of Anarchy*, which is the maximum ratio between the cost of a system optimal flow and the cost of a flow at equilibrium:

$$\rho(G, r, \ell) = \frac{\max_{f \text{ Nash}} C(f)}{\min_{f \text{ feasible}} C(f)} \ .$$

1.3 Characterising Optimal Flows

- Optimal flow: minimise the value of $C(f)$ among all feasible flows f.

- Formulation as a *convex program*:

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

subject to

$$\begin{array}{rcll} \sum_{P \in \mathcal{P}_i} f_P & = & r_i & \forall i \in [k] \\ f_e & = & \sum_{P \ni e} f_P & \forall e \in E \\ f_P & \geq & 0 & \forall P \in \mathcal{P} \end{array}$$

- The linear constraints of this program define a convex polyhedron.
- We assume that the terms $c_e(f) = f_e \cdot \ell_e(f)$ are convex.
- However, the number of variables is exponential in the network size (f_P). This can be reduced.

1.3.1 Marginal Cost

- What does shifting flow from path P_1 to P_2 imply?
- Cost of the edges in P_1 reduces, cost of P_2 increases.
- If cost of P_1 reduces more than cost of P_2 increases, total cost decreases, i. e., the flow is not optimal.
- Define *marginal cost*

$$\ell_e^*(x) = c_e'(x) = c_e(x)\frac{d}{dx}, \qquad c_P'(f) = \sum_{e \in P} c_e'(f_e) \enspace .$$

- Observation: A flow is optimal iff marginal cost of used paths is equal for all paths of the same commodity.

Lemma 2 (Characterisation of optima via marginal cost [8])**.** *A feasible flow f is optimal iff for all commodities $i \in [k]$, paths $P_1, P_2 \in [k]$, $f_{P_1} > 0$ we have $c_{P_1}'(f) \leq c_{P_2}'(f)$.*

Now, observe the similarity of this characterisation of optimal flows and Nash flows.

Theorem 3 (Equivalence of Nash equilibria and optima w. r. t. marginal cost)**.** *Assume $x \cdot \ell_e(x)$ is convex for all $e \in E$ and let $\ell_e^*(x) = (x \cdot \ell_e(x))\frac{d}{dx}$ for all $e \in E$. Then a feasible flow f is optimal for (G, r, ℓ) if and only if f is at a Nash equilibrium for (G, r, ℓ^*).*

Proof. The characterisations of optima for (G, r, ℓ) by Lemma 2 and Nash equilibria for (G, r, ℓ^*) by Lemma 1 are identical. □

Observe that $\ell_e^*(x) = (\ell_e(x)x)' = \ell_e(x) + \ell_e'(x)x$ consists of two terms. The first is the per-unit latency incurred by additional flow whereas the second accounts for the increased cost of the flow that is already using the edge.

1.4 Existence of Nash Equilibria

Theorem 4 (Existence and essential uniqueness of Nash equilibria [2])**.** *An instance (G, r, ℓ) with nondecreasing latency functions admits a flow at Nash equilibrium. Moreover, if f and $\tilde{f}$ are Nash equilibria, $C(f) = C(\tilde{f})$.*

Proof. We consider the following potential function due to Beckmann, McGuire, and Winsten:

$$H(f) = \sum_{e \in E} h_e(f_e) \quad \text{with} \quad h_e(x) = \int_0^x \ell_e(u)\, du \ .$$

Consider the convex program:

$$\min \sum_{i \in [k]} h_e(f)$$

subject to

$$\begin{array}{rcll} \sum_{P \in \mathcal{P}_i} f_P & = & r_i & \forall i \in [k] \\ f_e & = & \sum_{P \ni e} f_P & \forall e \in E \\ f_P & \geq & 0 & \forall P \in \mathcal{P} \end{array}$$

Proof of existence:

- Note that $h'_e(x) = \ell_e(x)$
- Hence, the optimality condition from Lemma 2 matches the characterisation for Nash equilibria from Lemma 1.
- Objective function H is convex (by monotonicity of the ℓ_e), and the solution space is nonempty and convex $\Rightarrow$ an optimum exists.

Proof of essential uniqueness:

- Consider Nash equilibria f and $\tilde{f}$
- Since both are feasible, the convex combination $f_\alpha = \alpha \cdot f + (1 - \alpha) \cdot \tilde{f}$ for $\alpha \in [0, 1]$ is also feasible by convexity of the solution space.
- h_e must be linear between f_e and $\tilde{f}_e$ since otherwise for some $\alpha \in (0, 1)$, we would have $H(f_\alpha) < H(f) = H(\tilde{f})$ which would contradict optimality of f and $\tilde{f}$.
- Hence, $\ell_e = h'_e$ must be constant between f_e and $\tilde{f}_e$ for all $e \in E$.
- Also $L_i(f) = L_i(\tilde{f})$ for all i and finally $C(f) = C(\tilde{f})$. □

1.5 Upper Bounding the Price of Anarchy

1.5.1 A Bound for Latency Functions of Limited Steepness

The price of anarchy depends on the steepness of the latency function. This is formalised by the following theorem.

Theorem 5. *Suppose for every $e \in E$ and all $x \in [0,1]$,*

$$x \cdot \ell_e(x) \leq \alpha \cdot \int_0^x \ell_e(u)\, du \ .$$

Then, the price of anarchy $\rho(G, r, \ell) \leq \alpha$.

Proof. Let f denote a Nash flow and f^* denote a system optimal flow for (G, r, ℓ).

$$\begin{aligned}
C(f) &= \sum_{e \in E} \ell_e(f_e) f_e && \text{(definition)} \\
&\leq \alpha \sum_{e \in E} \int_0^{f_e} \ell_e(u)\, du && \text{(hypothesis)} \\
&\leq \alpha \sum_{e \in E} \int_0^{f_e^*} \ell_e(u)\, du && \text{(since } f \text{ minimises } H\text{)} \\
&\leq \alpha \sum_{e \in E} \ell_e(f^*) f_e^* && \text{(by monotonicity of } \ell_e\text{)} \\
&= \alpha\, C(f^*) \ .
\end{aligned}$$

□

This yields an upper bound on the price of anarchy for polynomial latency functions:

Corollary 1. *Suppose latency functions have the form $\ell_e(x) = \sum_{i=0}^{d} a_{e,i} x^i$ for a positive integer d and $a_{e,i} \geq 0$. Then*

$$\rho(G, r, \ell) \leq d + 1 \ .$$

Proof.

$$x \cdot \ell_e(x) = \sum_{i=1}^{d} a_{e,i} x^{i+1} \leq (d+1) \sum_{i=1}^{d} \frac{a_{e,i}}{i+1} x^{i+1} = (d+1) \int_0^x \ell_e(u)\, du \ .$$

□

For linear latency functions, this yields an upper bound on the price of anarchy of two. A more involved analysis yields a better upper bound that matches our examples from the beginning:

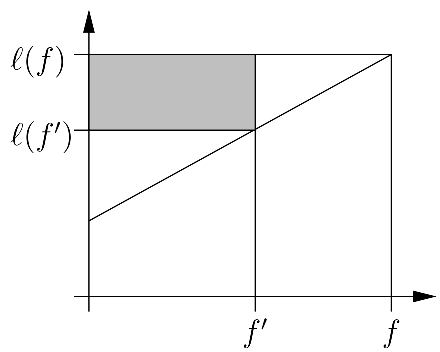
Theorem 6. *Suppose latency functions are linear with non-negative slope and offset. Then*

$$\rho(G, r, \ell) \le \frac{4}{3} \ .$$

Proof. Fix a Nash flow f and consider any other flow f'. Note that

$$\begin{aligned} C(f) &= \sum_{P \in \mathcal{P}} f_P \, \ell_P(f) \\ &\le \sum_{P \in \mathcal{P}} f'_P \, \ell_P(f) \\ &= \sum_{e \in E} f'_e \, \ell_e(f) \\ &= C(f') + \sum_{e \in E} f'_e \, (\ell_e(f) - \ell_e(f')) \end{aligned}$$

To see that the inequality holds consider the latency vector induced by f and note that f uses only paths with minimal latency whereas f' may (or may not) also use more expensive paths, but no cheper ones. The following figure shows that the last sum is at most $C(f')/4$. We consider only positive terms ($f'_e > f_e$), pessimistically ignoring the contribution of the negative terms.



Obviously, the upper left rectangle of size $f'_e \, (\ell_e(f) - \ell_e(f'))$ is at most one fourth of the entire rectangle of size $f_e \, \ell_e(f)$.

Since our inequality holds for all f' and, in particular, for an optimal flow, this proves the theorem. $\square$

1.5.2 A Bicriteria Bound

Theorem 7. *If f is a flow at Nash equilibrium for (G, r, ℓ) and f^* is feasible for $(G, 2r, \ell)$, then $C(f) \le C(f^*)$.*

Proof.

- We construct a helper instance with latency functions where the cheap part is cut away:

$$\bar{\ell}_e(x) = \begin{cases} \ell_e(f_e) & \text{if } x \le f_e \\ \ell_e(x) & \text{if } x \ge f_e \end{cases}$$

- Denote the respective cost function by

$$\bar{C}(f) = \sum_{e \in E} f_e \cdot \bar{\ell}_e(f_e)$$

- We bound the cost we can "lose" by using these latency functions. For any edge $e \in E$ and $x \geq 0$:

$$x(\bar{\ell}_e(x) - \ell_e(x)) \leq \ell_e(f_e) f_e$$

(If $x \geq f_e$ the difference is zero, and if $x \leq f_e$, the difference is maximised to $\ell_e(f_e)f_e$ if ℓ_e drops to 0 left of f_e.)

- Now consider the total additional cost by evaluating f^* with respect to $\bar{C}$ instead of C:

$$\bar{C}(f^*) - C(f^*) = \sum_{e \in E} f_e^*(\bar{\ell}_e(f_e^*) - \ell_e(f^*)) \leq \sum_{e \in E} \ell_e(f_e) f_e = C(f) \tag{1}$$

- Let 0 denote the zero flow. Then, for commodity $i \in [k]$ and $P \in \mathcal{P}_i$

$$\bar{\ell}_P(f^*) \geq \bar{\ell}_P(0) \geq L_i(f)$$

- Hence,

$$\bar{C}(f^*) \geq \sum_{i \in [k]} \sum_{P \in \mathcal{P}_i} L_i(f) f_p^* = \sum_{i \in [k]} 2L_i(f) r_i = 2C(f) \tag{2}$$

- Combining Equations (1) and (2) we obtain

$$C(f^*) \geq \bar{C}(f^*) - C(f) \geq 2C(f) - C(f) = C(f) \qquad \square$$

2 Congestion Games

This section presents some of the results of [1, 4, 6]. The existence proof of Nash equilibria in congestion games can be found in [6], the reduction from Min-Cost Flow can be found in [4], and the PLS-hardness result can be found in [1].

2.1 The Model

- Very similar to Wardrop Model
- First difference: Finite number of agents n, each controlling unit load
- Second difference: Generalised combinatorial structure of strategy space: edge set replaced by set of abstract *resources* E, the set S_i of strategies for a player i is no longer necessarily a set of paths, but is specified explicitly as a set of subsets of E.

- For any player $i \in [n]$ let $s_i \in S_i$ denote the strategy chosen by player i.
- For $e \in E$, the load (flow) of resource e is now the number of agents selecting a strategy containing this resource: $f_e(s) = |\{i \mid e \in s_i\}|$.
- A congestion game is *symmetric* if for any $i, j \in [n]$, $S_i = S_j$.
- A congestion game is a *network congestion game*, of the resources are edges of a network and the strategies S_i are sets of paths connecting a sink and source as in the previous section.
- Rest of definitions unchanged.
- A *strategy profile* $s = (s_1, \ldots, s_n) \in S_1 \times \cdots \times S_n$ is at a *Nash equilibrium* if for any i and any $s'_i \in S_i$ it holds that $\ell_{s'_i}(s_1, \ldots, s'_i, \ldots, s_n) \geq \ell_{s_i}(s_1, \ldots, s_i, \ldots, s_n)$.

2.2 Existence of a Nash equilibrium

- Potential function H can be used in a discrete version Φ [6]:

$$\Phi(s) = \sum_{e \in E} \sum_{k=1}^{f_e(s)} \ell_e(k)$$

- Fix a strategy profile s and a strategy s' which differs from s only in the strategy played by some player i. Then, the terms in the outer sum of Φ differ only for those resources that are contained in the symmetric difference of s_i and s'_i.

$$\begin{aligned} \Phi(s') - \Phi(s) &= \sum_{e \in s'_i \setminus s_i} \ell_e(f(s')) - \sum_{e \in s_i \setminus s'_i} \ell_e(f(s)) \\ &= \ell_{s'_i}(s') - \ell_{s_i}(s) \end{aligned}$$

- Note that this is precisely the improvement made by the defecting player i.
- Hence, no player can make an improvement iff $\Phi(s)$ is at a minimum. Since Φ clearly attains a minimum over the finite strategy space, we have proven:

Theorem 8. *Every congestion game has a pure Nash equilibrium.*

Note that this also yields an (exponential) algorithm for computing a Nash equilibrium: Start with an aribtrary strategy profile and let the players perform improvement steps until no more improvement is possible.

2.3 Symmetric Network Congestion Games

For symmetric network congestion games, it is possible to compute Nash equilibria in polynomial time by a reduction to min-cost flow:

- Take as input a network $G = (V, E)$, latency functions $(\ell_e)_{e \in E}$, a source s and a sink t
- Replace each edge $e \in E$ by n parallel edges $e_1, \ldots, e_n$, each with capacity 1. Edge e_j has cost $\ell_e(j)$
- Compute an integer min-cost flow f for this network.
- Clearly, a min-cost flow f utilising k of the edges $e_1, \ldots, e_n$ will use precisely the edges $e_1, \ldots, e_k$ and contribute cost $\sum_{j=1}^{k} \ell_e(j)$.
- Hence, a min-cost flow corresponds to a flow in G that minimises Φ.

2.4 The Asymmetric Case

- In the asymmetric case, this reduction is not possible
- To show hardness, we consider the complexity class PLS.
- A search problem Π in PLS is given by a set of instances $\mathcal{I}_\Pi$ and for every $I \in \mathcal{I}_\Pi$
  1. a feasible set $F(I)$, an objective function $c : F(I) \to \mathbb{N}$, and a neighbourhood $N(S, I) \subseteq F(I)$ for all feasible solutions $S \in F(I)$.
  2. a polynomial time algorithm A that computes for every $I \in \Pi$ an initial feasible solution $S^0 \in F(I)$.
  3. a polynomial time algorithm B that computes for every $I \in \Pi$ and $S \in F(I)$ the objective value $c(S)$
  4. a polynomial time algorithm C that determines for every $I \in \Pi$ and $S \in F(I)$ whether S is locally optimal and, if not, finds a better solutin $S' \in N(S, I)$.
- An instance of the PLS problem is to find a local optimum, i. e., find a feasible solution S such that algorithm C returns **“yes”**.
- PLS contains many hard problems like search versions of weighted satisfiability and TSP. No polynomial algorithms are known.
- A reduction from a PLS-problem Π_1 to a PLS-problem P_2 is given by two functions f and g computable in polynomial time such that
  1. f maps instances I of Π_1 to instances $f(I)$ of Π_2,
  2. g maps pairs (S', I) where S' is a solution of $f(I)$ to soultions of I,
  3. for all $I \in \Pi_1$, solutions S' of $f(I)$, if S' is locally optimal for $f(I)$, then $g(S', I)$ is locally optimal for I.

2.4.1 Max-Cut

- Input: A complete graph $G = (V, E)$ with edge weights $(w_e)_{e \in E}$
- Search space: Set of cuts (a pair (U, W) such that U and W are disjoint and $U \cup W = V$.
- Neighbourhood: Two cuts are neighboured, if they differ only in one vertex.
- Objective function: Sum of weights of edges crossing the cut.
- It is PLS-hard to find a local optimum for Max-Cut.
- Can be viewed as a *party affiliation game*. Every node chooses one side of the cut. It prefers the side where the edge weights are at least one half of the overall weight of incident edges.

2.4.2 Quadratic Threshold Games

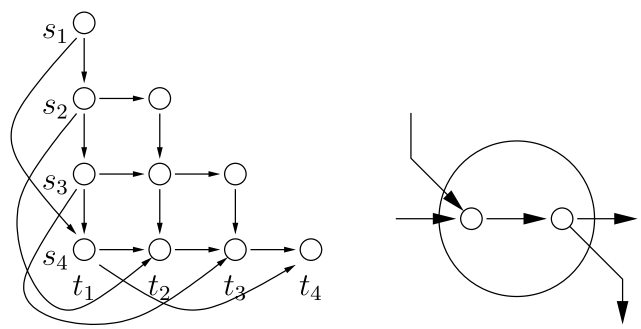
- Special class of congestion games where each player has only two strategies.
- Two classes of resources $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$
- First strategy $S_i^{\text{out}} = \{r_i\}$, $r_i \in \mathcal{R}_{\text{out}}$ where r_i has a fixed delay T_i, the *threshold* of player i.
- Second strategy: $S_i^{\text{in}} \subseteq \mathcal{R}_{\text{in}}$.
- For Quadratic Threshold Games (QTG), $\mathcal{R}_{\text{in}}$ contains exactly one resource r_{ij} for every unordered pair of players $\{i, j\}$ and $S_i^{\text{in}} = \{r_{ij} \mid j \in [n],\ j \neq i\}$.

Theorem 9. *Computing a Nash equilibrium for Quadratric Threshold Games is PLS-complete.*

Proof. Reduction from Max Cut.

- Set of nodes in Max Cut becomes set of players in QTG.
- For every edge $e = (i, j)$, introduce resource $r_{ij} \in \mathcal{R}_{\text{in}}$; its delay function is 0 if used by one player and w_{ij} is used by two.
- The delay of the threshold resource $r_i \in \mathcal{R}_{\text{out}}$ of player i is $T_i = \frac{1}{2} \sum_j w_{ij}$.
- Game is isomorphic to party affiliation game.

$\square$



(a) Construction for $n = 4$ players.

(b) Gadget for nodes with delay functions.

Figure 1: Illustrations for the reduction from QTG to Network Congestion Games.

2.4.3 Network Congestion Games

Theorem 10. *Computing a Nash equilibrium for Network Congestion Games is PLS-complete.*

Proof. Reduction from QTG.

- Network is lower left triangle of an $n \times n$-grid. Edges are directed right and downward, respectively.
- Player i's source node s_i is the i-th node in the first column (counted from top) and her target node t_i is the $i-th$ node of the last row (counted from left).
- In addition, add *threshold edge* (s_i, t_i).
- See Figure 1(a).
- Now, design delay functions such that there are only two relevant strategies: the threshold edge, and the *row-column-path* which first goes only right, and then only down.
  - Fix large integer $D > \sum_{i,j} w_{ij}$.
  - Threshold edge has constant delay $D \cdot i \cdot (i-1) + T_i$
  - Column edges have constant delay 0
  - Row edges have delay $D \cdot i$.
  - Also define delay functions for the nodes (and replace them by a gadget consisting of an in-node connected to incoming edges, an out-node connected to outgoing edges, and one edge connecting in-node and out-node which is actually assigned the delay functions, see Figure 1(b))

– Delay function of node in column i and row j is delay function of resource r_{ij} from QTG.

- Now, the threshold edge corresponds to S_i^{out} and the row-column-path corresponds to S_i^{in} (both with an additional cost of $D \cdot i \cdot (i-1)$. All other strategies are dominated since using an edge from a row below would cost an additional D.

- Hence, Nash equilibria of both games coincide.

□

References

[1] Heiner Ackermann, Heiko Röglin, and Berthold Vöcking. On the impact of combinatorial structure on congestion games. In *Proc. 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 613–622, 2006.

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

[3] José R. Correa, Andreas S. Schulz, and Nicolás E. Stier-Moses. On the inefficiency of equilibria in nonatomic congestion games. In *Proc. 11th International Integer Programming and Combinatorial Optimization Conf. (IPCO)*, volume 2509 of *Lecture Notes in Computer Science*, pages 167–181. Springer, 2005.

[4] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure Nash equilibria. In *Proc. 36th Annual ACM Symposium on Theory of Computing (STOC)*, pages 604–612, 2004.

[5] Alain B. Haurie and Patrice Marcotte. On the relationship between Nash-Cournot and Wardrop equilibria. *Networks*, 15:295–308, 1985.

[6] Robert W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[7] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[8] John Glen Wardrop. Some theoretical aspects of road traffic research. In *Proc. of the Institute of Civil Engineers, Pt. II*, volume 1, pages 325–378, 1952.