Topics in Artificial Intelligence: Multiagent Systems
Selfish Routing in Computer Networks

Sebastian Streg

December 10, 2005

1 Introduction

Efficiency in networks with lots of traffic is a serious problem in many different fields. Road traffic for example had been studied for several decades. In computer networks like the Internet this problem has also become much more important over the past few years because of the growing amount of data. No matter what kind of networks you study, they share many properties if you abstract the model. All networks have nodes, edges and traffic had to be routed through the network as efficiently as possible. So you can transfer many results from other, where network problems had been studied to computer networks.
Network efficiency is hard to describe. Network users usually want to route their traffic at lowest cost. Costs might be waiting time or taxes for using edges. From a central point of view you have two ways of looking at efficiency. You can route traffic in that manner that the sum of costs from each edge is minimal. Or you might want the maximum cost of a user to be as low as possible. For example you have to route two files at the same time through the network. You have the possibility that both files need 30 seconds or one needs 10 seconds and the other one needs 40 seconds. In the first time you need more resources because the sum is higher but you are done 10 seconds faster.
This paper deals with the problem of efficient routing in networks with selfish routers without regulation. I want to analyze if the outcome of network with a central regulation is worse than one without. This work will show that there exists a trade off between that both possibilities. Furthermore, I will present possibilities to achieve the optimal outcome with the help of taxation and how powerful they are. After that I will discuss the inequity some users may experience if network designers try to achieve a optimal outcome.

2 Model

First of all we have to introduce some Variables which describe our network. Our network is a graph (V, E) which consists of nodes and edges. Traffic should be carried from one node to another. This traffic is our flow f. Every flow has a source s and a target t. There are networks which have only one unique source and one unique target whereas in some networks every node can be target or source, or both. The results presented in this paper consider the general case, but some examples use the simple case to provide a better illustration.
In the Introduction we said that we want users to minimize costs while using the network. Costs can be expressed in terms of latency. Every edge e of a network has its own latency $l_e(f_e)$ which is an increasing function of the flow of this edge. Consequently the latency of a Path P is

- $l_p(f) = \sum_{e \in E} l_e(f_e)$.

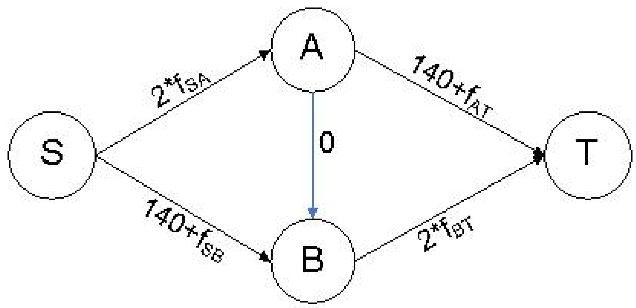

figure1

As the cost of all flows or total latency we define as:

- $C(f) = \sum_{P \in P} l_P(f) f_P = \sum_{e \in E} l_e(f_e) f_e$

A Wardrop Model is a network model that judges an outcome by the total latency or the average latency per user. Both approaches are equivalent. In contrast to that a KP-Model[1] looks at the maximum latency of a network user.
Minimizing the total latency means minimizing social costs:

- $MinC(f) = Min(\sum_{e \in E} l_e(f_e) f_e) = Min(\sum_{e \in E} c_e(f_e))$

A flow which minimizes social costs is called optimal flow. A Nash equilibrium is a set of strategies of each player such as that no player has incentive to unilaterally change his action like in the famous "Prisoner's Dilemma". [2]Nash equilibria do not necessarily minimize social costs. A flow is in Nash equilibrium if no agent can improve its latency by changing its path. We will call that flow a Nash flow.

3 Price of Anarchy

In a network with a lack of a central regulation users are non-cooperative. That means that they act selfishly. The user's aim is to route their traffic at lowest cost through the network. However the shortest path is not always to best one. Imagine you have two mirrors where you can download a file. One has a total bit rate of 10 Mbit/s the other one 1 Mbit/s. When 20 internet users want to download the same file, all of them should choose the first mirror. But each user will have an average download rate of 0,5 Mbit/s. So one user would have incentive to use the second mirror if he was the only one.
Brass' Paradoxon[3] becomes very popular in the network theory, because it proves that there are networks with selfish routers where a removal of edge can improve the efficiency of an outcome. The first article about that topic was already published in 1968 while analyzing road traffic. Let's consider figure 1. To keep it simple we have a network with one starting point and one target. The edge latencies are given in the graph. There are 3 paths to route your traffic from S to T (P_{SAT}, P_{SABT} and P_{SBT}). The latencies for each way are:

[1]see [13] to learn all different aspect of the Wardrop and KP Model.
[2]see [10],p.80f
[3]see [8]

- $l_{SAT} = 140 + 2 * f_{SA} + f_{AT}$
- $l_{SBT} = 140 + 2 * f_{BT} + f_{SB}$
- $l_{SABT} = 2 * f_{SA} + f_{BT}$

F is the flow which has to be routed through the network. Let's assume that each user wants to route exactly a flow of $f = 1$. If we had 20 users it is obvious that each of them will use P_{SABT}, since the total latency is 80 and not as big as the constant latency factor of the other ways. Now suppose we have 100 users which want to route their traffic through the network. All traffic has to be routed from S to T. So we have a total flow $f = 100$. If we had a central regulation and not selfish routers, this regulation would determine that 50 users will use P_{SAT} and 50 users P_{SBT}, so that all of them experience a latency of exactly 290. Note that the latency of the way P_{SABT} is 200 at the moment that nobody uses it. But when we have selfish routers this chosen flow is an optimal flow, but not a Nash flow. A actor which uses P_{SAT} could minimize his own latency by deviating to P_{SABT}. He would have a latency of 202 whereas a P_{SAT} user has now a latency of 289, but P_{SBT}'s latency increase to 292. Of course a P_{SBT} user has the same incentive to use P_{SABT}, increasing P_{SABT}'s latency to 204, decreasing P_{SBT}'s latency to 291 and increasing P_{SAT}'s latency to 291. We see that when one actor using P_{SAT} and one actor using P_{SBT} defect by deviating toward P_{SABT}, l_{SAT} and l_{SBT} increases 1 units, but l_{SABT} increases 4 units so that after 60 routers had defected we can find a Nash equilibrium. 20 users are using P_{SAT}, 20 user's P_{SBT} and 60 users P_{SABT} and all of them had a latency of 320. Now no user has incentive to deviate because he would increase the latency of the way to which he deviates. So the outcome is much worse compared to the outcome made by a central regulation to everyone. However, if we removed edge AB from our graph, the way P_{SABT} would become impossible and nobody could deviate towards P_{SABT}. The optimal flow would become a Nash flow. So we see that removing the most useful edge(because it has a constant latency of 0), makes our network more efficient.

Braess' Paradoxon stated that in a network an optimal flow is not necessarily a Nash flow. However, we can try to find a upper bound for the worst flow. Therefore, there exists a Coordination ratio, which is well known as "the price of anarchy"' in selfish networks. It is defined: $\rho = \frac{C(f)}{C(f*)}$. Roughgarden and Tardes[4] showed that the coordination ration in the worst-case Nash equilibrium is not more than 4/3 under the assumption of having linear latency functions. Furthermore they showed that for increasing latency functions, which are polynomials with degree p higher than 2, the worst coordination ratio for the worst case Nash equilibrium is $\theta(\frac{p}{logp})$

4 Edge Taxation

In the previous section we talked about the efficiency of Nash equilibria. We stated that there exists at least one pure strategy Nash equilibrium in networks with selfish users. Since it is not necessarily unique, there might be more than one Nash equilibrium and we described the worst-case Nash equilibrium and introduced the coordination ratio.
However there exists a discussion in the recent research work on how to design networks to make Nash equilibria more efficient or in the best case how optimal flows will become unique Nash flows. We will discuss three possibilities[5], how you might influence optimal flow and Nash flow: edge removal, arbitrary taxation and marginal cost taxation.

[4]see [7]
[5]all claims in that paragraph are based on[7]

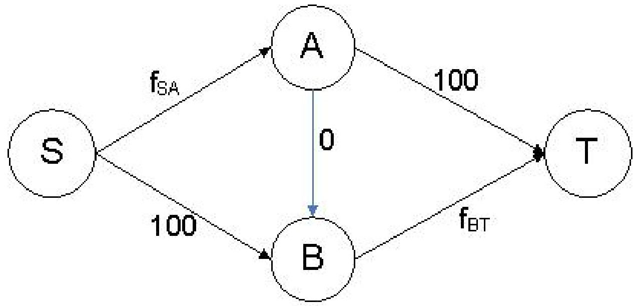

figure 2

The first approach we have already described in the last section. Brass' Paradoxon showed that edge removal can achieve a Nash equilibrium which optimizes social costs. So as a network designer you act as a central regulation under game theoretical aspects. You have to recognize inefficient Nash equilibria and remove edges in that way that you cannot constitute these equilibria anymore. You try to achieve through edge removal that the optimal flow will become a Nash flow. In our example of the Braess Paradoxon it would mean to remove edge (A, B). However, not every network allows you to achieve the social optimum. Especially in large networks with non-linear latency-functions the computation of eliminating all "bad" Nash flows can be achieved only in exponential time. Sometimes after edge removal the social optimum is still not a Nash equilibrium and sometimes it is, but the social costs of that optimal flow are higher than before edge removal. A further problem might be that removing edges means destroying infrastructure, especially in cases where you use the same network in different manners. For example you have different flows or different source and target nodes over time, you will have different optimal flows using different edges.
Cole, Dodis and Roughgarden also analyzed an equivalent method: arbitrary taxes. That means that users pay for using edges to route their traffic and the price is calculated independently of the latency and the delay which a user causes to the others by using the network. In our Braess' Paradoxon example we would set the tax for (A, B) very high, say 2000 units, and no tax for the other edges so that nobody has the incentive to use Path P_{SABT} anymore. We would have the same effect like removing those edge from the network and everybody would head to the social optimum outcome which is a Nash flow after introducing that tax. However, Cole, Dodis and Roughgarden proved that removing that arbitrary taxes cannot improve efficiency better than edge removal under the assumption having linear latency functions. But under the assumption of having polynomial latency function with degree greater than 2 this equivalence statement does not hold, because the improvement might be better than edge removal. The proof can also be seen in[7]. In contrast to edge removal, taxation seems to be better to handle in practice, because you can change the amount of taxes which users have to pay for an edge very quickly.
Another proposed way to achieve a better flow in Nash equilibrium might be marginal cost taxation, meaning that a tax for using an edge of a network is higher, when the user causes a higher delay to other users. Let's consider figure 2 which is more simple a Braess'Paradoxon example. We want to compute optimal marginal taxes. The optimal flow is $f_{SAT} = 50$ and $f_{SAT} = 50$ whereas the Nash flow is $f_{SABT} = 100$. By using edge (S, A) a user causes 1 unit more harm to the others than when he uses (S, B). The same applies to (B, T) compared to (A, T). So we have to add a tax t to (S, A) and (B, T) of $t = 1 * f$. Then the optimal flow is a Nash flow, but the prevoius Nash equilibrium is still a Nash equilibrium.

But compared to the old optimal flow the costs has changed. The costs are now 200 and so we can say that $S_{opt,new} = \frac{4}{3} S_{opt,old}$ so that the social optimum is worse than before. So we did not

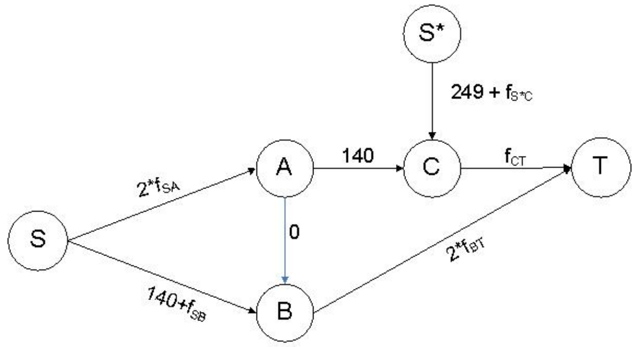

figure 3

improved costs of our Nash flow and achieved no improvement in efficiency. We have just damaged the optimal flow. Efficiency could only be achieved when taxes would be refunded, for example by paying the same amount to each user.[6] One thing that the authors did not consider in their models is tax sensitivity. Each user reacts different to taxes, because comparing time delay to money, it might have a different value to him. Consider for example prices for flight ticket. You can choose between a direct flight or a flight with 2 stops which is 20 percent cheaper. For a business man saving time is more important than spending a little bit more money so he will probably choose a direct flight, whereas a student might decide to buy the cheaper flight. This sensitivity had not been considered when looking at taxation in networks. [7]

5 Fairness of optimal flows

We used the Wardrop model for routing in networks so for as it is the most common one to use, because it considers average total latency. We tried to find a way to to make Nash flows better. However, heading to the social optimum burdens a problem which is hardly discussed in the literature so far: equality. There might be users in a social optimum which might be penalized harshly. Let's consider following example. You will recognize that figure 3 is also a Braess' Paradoxon example, but we did again some modifications. Say again we have 100 users where each of them wants to route a flow f=1 from s to t, so that the total flow is 100 so far. However we have an additional user, user U_{101}. This user is highly sensitive to latency. Since he does not want to be dependent to others and their decisions on how to use the network he invests much into a new edge to connect to the network. U_{101} is able to have a different source S^* from which he can reach T much faster. U_{101} supposes that network users act selfishly, and that nobody wants to introduce taxation or edge removal. Since he is rational and thinks that the others are supposed to be rational as well, he knows that all the other users will constitute the same Nash flow like in section 3. This means that 60 users will use P_{SABT}, 20 users P_{SACT} and another 20 users P_{SBT}. Note that the existence of U_{101} does not change anything. For the P_{SACT} users he causes increase the latency of one unit, but they cannot improve by deviating. However, this Nash flow is not unique anymore, because exactly one P_{SACT} can switch to either P_{SABT} or P_{SBT}.
What happened now when we try to achieve optimal social costs by techniques proposed in section 4? Among the 100 remaining users, 50 will use P_{SACT} and the other 50 P_{SBT}. These users will

[6] see[7], p.4

[7] Karakostas and Kolliopoilos[12] considered that problem in their research work. Their results are not reviewed in that paper.

have a latency of 290 compared to 320 before. But U_{101} will have 300 instead of 270. He is now worse off using his own way than sharing the edges with the others. He is really penalized by the virtual central regulation in form of the network designers which want to improve Nash flows through taxes or edge removal. The main reason why the example in section 3 does not always work (truly said, it will work in a very few cases) is that the Nash flow was pareto dominated by the optimal flow. The same phenomena was detected in the prisoner's dilemma, where the outcome constituted by the only Nash equilibrium was the only outcome which was not pareto efficient. In that example the nash flow is not pareto dominated because we saw that U_{101} is now worse off.

6 Summary

This paper reflects many current research results about routing in networks with selfish users. It presents that optimal flows in networks are not necessarily Nash flows and that in the case of Braess' Paradoxon a Nash flow might even be pareto dominated by a optimal flow. It also presents possible solutions to network designers to make Nash flows more efficient, like edge removal or taxation. In theory, there might be a way out of the dilemma, but we saw that there are barrier like complexity and heterogeneous users that network designers have to consider. At the end the paper showed that this sort of central regulation can discriminate some users while trying to achieve the optimal flow. We saw that coping with networks is very hard and we could not find a solution which is satisfactory to every aspect network designers have to consider.

7 References

1. T. Roughgarden. *Selfish Routing and the Price of Anarchy*,MIT Press, May 2005

2. T. Roughgarden. *The Maximum Latency of Selfish Routing*,SODA 2004, pages 973–974

3. T. Roughgarden and E. Tardos, *Bounding the Inefficiency of Equilibria in Nonatomic Congestion Games, Games and Economic Behavior*, 47(2):389–403, May 2004

4. T. Roughgarden. *The Price of Anarchy is Independent of the Network Topology*, STOC 2002, pages 428–437

5. T. Roughgarden and E. Tardos. *How Bad is Selfish Routing?*, FOCS 2000, pages 93–102.

6. R. Cole, Y. Dodis, T. Roughgarden. *Pricing Networks with Selfish Routing*, STOC 2003, pages 521-530.

7. R. Cole, Y. Dodis, T. Roughgarden. *How Much Can Taxes Help Selfish Routing?* , EC 2003, pages 98–107.

8. D.Braess. *Ueber ein Paradoxon der Verkehrsplanung*, Unternehmensforschung, 12:258-268, 1968

9. E. Koutsoupias, C. H. Papadimitriou. *Worst-case equilibria*, STACS 99

10. Yoav Shoham, Kevin Leyton-Brown, *Multiagent systems*, unpublished

11. Cem U. Saraydar, Narayan B. Mandayam, David J.Goodman: *Efficent Power Control via Pricing in Wireless Networks*,IEEE Journal on Selected Areas in Communications 19(10): 1883-1892 (2001)

12. Karakostas and Kolliopoilos, *Edge Pricing of Multicommodity Newtowrks for Heterogenous Selfsh Users*

13. Rainer Feldmann, Martin Gairing, Thomas Lcking, Burkhard Monien, Manuel Rode. *Selfish Routing in Non-Cooperative Networks: A Survey*, Proceedings of the 28th International Symposium Mathematical Foundations of Computer Science (Branislav Rovan, P. Vojtas eds.) Springer Lecture Notes in Computer Science (LNCS 2747), pp 21-45, 2003.