A Competitive Strategy for Routing Flow over Time

UMANG BHASKAR
Dartmouth College
and
LISA FLEISCHER
Dartmouth College
and
ELLIOT ANSHELEVICH
Rensselaer Polytechnic Institute

Network routing games are used to understand the impact of individual users' decisions on network efficiency. Prior work on routing games uses a simplified model of network flow where all flow exists simultaneously. In our work, we examine routing games in a flow-over-time model. We show that by reducing network capacity judiciously, the network owner can ensure that the equilibrium is no worse than a small constant times the optimal in the original network, for two natural measures of optimality. These are the first upper bounds on the price of anarchy in the flow-over-time model for general networks.

1. INTRODUCTION

We are interested in understanding the behavior of traffic in large networks, such as the Internet and road networks. There are two common features of traffic in these networks. (1) Traffic travels over time. Thus, the state of these networks is a function of time, and the shortest-path through the networks varies over time. (2) The traffic is not centrally coordinated. Units of traffic are routed to optimize some local and individual objective, such as the time of arrival at the destination.

Routing games are a widely used tool to model the behavior of traffic in networks. In these games, players route traffic in a network with the objective of minimizing their cost. Thus, routing games capture the decentralized nature of traffic routing. In particular, the price of anarchy — the ratio of cost of the worst equilibria to the cost of the optimal routing — has emerged as a standard for measuring the inefficiency caused by the lack of coordination among players. Much of the popularity of routing games owes to the clarity of the models which afford rigorous analysis, and to the tight bounds obtained on the price of anarchy.

Most previous studies of routing games model flows as static: the flow on every edge is fixed, and occurs instantaneously and simultaneously. In contrast, many

commercially used packages for traffic analysis, in order to capture the dynamic nature of traffic, do not make this assumption, e.g., DYNASMART [Mahmassani et al.]. These packages view traffic as inherently time-varying. The price paid for the added complexity is that these models cannot often be rigorously analysed, and instead approaches such as simulation are used.

We study a deterministic queueing-based model for routing games based on flows-over-time. In this model, each edge in the network has a fixed capacity that limits the rate of flow through the edge. Flow entering an edge at a rate in excess of the edge capacity waits in a queue, and clears the queue in FIFO order. Each edge also has an edge-delay, which is the time taken by a flow particle to traverse the edge once it clears the queue. Thus, excess flow entering an edge causes congestion in the form of increased queues, which increases the time taken to traverse the edge. Traffic in this model travels through the network over time, and the state of the network, including the congestion on edges, is a function of time.

As an initial step, we consider the case when all traffic travels from a single source to a single destination, and players are nonatomic, i.e., each player controls infinitesimal flow. Each player picks as a strategy a route from the source to the destination to minimize the arrival time of her flow at the destination. This choice differs for different players, since the shortest path through the network varies with time. We refer to routing games in this model as *temporal routing games*. In addition to modeling internet and road traffic, temporal routing games can also be used to model evacuation traffic. Equilibria exist in these games [Cominetti et al. 2011; Koch and Skutella 2008].

This model has a number of advantages over traditional models. Previous models with static flow assumptions measure the cost of a player as either the total delay or the maximum delay on any edge used by the player, depending on the application. The total delay is used to measure a player's cost when the travel time is important, while the maximum delay is used when the throughput is important. Both of these measures (travel time and throughput) are substitutes for measuring the time by which all traffic arrives at the destination, and the use of these substitutes often yields counterintuitive examples. E.g., in bottleneck congestion games, which seek to model applications where the throughput is important, a player's cost is the delay of the maximum delay edge used by the player, and the delay on other edges used by a player are ignored. This leads to examples where players use a large number of system resources when a smaller set of resources would suffice, since they only pay for the use of the most congested resource. A number of examples with large price of anarchy are based on this counterintuitive behavior by the players. The price of anarchy improves if users are charged for usage of edges besides the maximum delay edge [Cole et al. 2006].

Routing games with static flows further require a delay function to be defined on every edge, which gives the delay on the edge as a function of the flow on the edge. The performance of equilibrium is critically dependent on these functions, which for a network may be difficult to obtain. For example, for nonatomic routing games with polynomial delay functions, the price of anarchy is $\Theta(\frac{d}{\log d})$, where d is the maximum degree of the delay function on any edge [Roughgarden 2003].

In contrast, in temporal routing games, the time at which traffic arrives at the

destination can be measured directly, and hence we do not need to use substitute measures. The cost of a player in a temporal routing game is simply the time she arrives at the destination. Also, each edge in a temporal routing game has a fixed capacity and edge-delay, and congestion on edges is represented through queues. Thus, the total delay depends on the traffic through the queueing delays, and we do not need to be given delay functions on the edges.

2.   THE STRUCTURE OF EQUILIBRIA

The equilibrium flow in temporal routing games can be described via a sequence of events [Koch and Skutella 2009]. An *event* corresponds to either a new edge in the graph being used by players, or a queue on an edge emptying. A *phase* is the time between successive events. Within a phase, the equilibrium flow can be described in terms of a special static flow, termed a *thin flow*.

For an edge $e = (v, w)$, consider the flow which leaves the source at time θ, and let $l_v(\theta)$ be the time this flow arrives at vertex v. Let $F_e^+(l_v(\theta))$ be the total flow which has entered edge e by time $l_v(\theta)$. Then the thin flow on edge e is $\frac{\partial F_e^+(l_v(\theta))}{\partial \theta}$. The thin flow on an edge thus corresponds to the product of two terms, where the first term $\frac{\partial F_e^+(l_v(\theta))}{\partial l_v(\theta)}$ is the rate of flow entering the edge, and the second term $\frac{\partial l_v(\theta)}{\partial \theta}$ is the rate of change of arrival time at the edge for flow which leaves the source at time θ. The thin flow on an edge is constant within a phase.

Given an algorithm for computing thin flows, the equilibrium flow can be computed using phases. Within a phase, we compute the thin flow, and use it to obtain the flow-over-time and queue on each edge. We keep track of queue sizes and edges used by players. Whenever a new edge starts to be used, or a queue empties, we end that phase, recompute the thin flow with the new edges and queues, and start a new phase. This continues until all the flow has reached the sink.
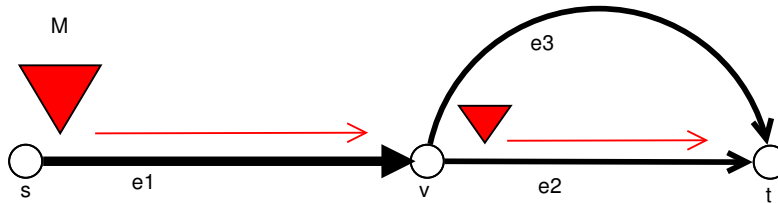

Fig. 1. A temporal routing game. Edge 1 is (s, v), edge 2 is the lower (v, t) edge, and edge 3 is the upper (v, t) edge. The capacity and edge-delay on edge i are c_i and d_i respectively. The inverted triangles indicate the FIFO queues on the edges in phase 2.

Figure 1 shows a simple temporal routing game. M units of flow at the source s are trying to reach the destination t. In the first phase of the equilibrium flow, each player uses only the bottom path, since each edge on this path has 0 edge-delay. Edge 2 has lower capacity than edge 1, hence a queue grows on edge 2 (depicted in the figure). Due to the increasing queue size, successive players arriving at edge 2 will face greater delays. When the time to be spent waiting in the queue on edge 2 is the same as the edge-delay on edge 3, players will start to use edge 3 in addition to

edge 2. Thus, in the second phase, all three edges are used by the players. Players arriving at vertex v will now split evenly between edges 2 and 3, and this continues until M units of flow reach t.

For comparison, in the centrally-coordinated flow-over-time that minimizes the time for the M units to reach the destination, edge 3 would be used concurrently with edge 2 from the beginning, and no queue would form. At equilibrium, due to the queue on edge 2 at equilibrium, any player using edge 2 must wait in queue before she traverses the edge. The queue thus increases the travel time for all players using edge 2. Queues are hence responsible for inefficiencies in temporal routing games.

3.   PRICE OF ANARCHY IN TEMPORAL ROUTING GAMES

The structure of equilibria in routing games allows us to obtain bounds on the worst-case behavior of equilibrium, with respect to a definition of social cost. This is formalized as the *price of anarchy*, defined as the ratio of the social cost at the worst equilibrium to the minimum social cost achievable under centralized control. For example, in temporal routing games, a measure of social cost is the amount of flow that has reached the sink at any given time. This is particularly relevant in the case of an imminent disaster, e.g., a building collapse. The price of anarchy defined for this measure is called the *evacuation price of anarchy*, and can be as large as $\Omega(m)$, where m is the number of edges in the graph [Koch and Skutella 2009]. A different measure is the maximum travel time — the maximum time taken by any player to travel from the source to the sink. Here, too, the price of anarchy is known to be as large as $\Omega(m)$ [Macko et al. 2010].

We study the price of anarchy with respect to two other natural measures of social cost: (1) the time when the last player reaches the sink, and (2) the average time of arrival of the players at the sink [Bhaskar et al. 2011]. We call the price of anarchy defined for the first measure the *time price of anarchy*, and for the second measure the *total delay price of anarchy*. In contrast to the results mentioned above, we give a simple mechanism for reducing capacities in the network that ensures that the price of anarchy for both measures we study is $O(1)$ with respect to the optimal flow for both the new capacities and original capacities. The constants we obtain are small, and in the case of the first measure, the bound obtained is tight.

The quickest flow is the flow-over-time which minimizes the time when the last player reaches the sink, and it can be computed in polynomial time [Ford and Fulkerson 1962]. Our mechanism proceeds by computing the quickest flow in the given network, and reducing the capacity of each edge to that used by the quickest flow. We show that in a network where the quickest flow uses the entire capacity of each edge, the time price of anarchy is bounded by $e/(e-1)$, while the total delay price of anarchy is bounded by $2e/(e-1)$. Thus in the modified network, both the time price of anarchy and total delay price of anarchy are bounded by these constants. Since the reduction in edge capacities does not affect the quickest flow, the bound on the time price of anarchy is with respect to the quickest flow in the original network. We show that the bound on the total delay price of anarchy is with respect to the optimal flow in the original network as well.

Our analysis for the time price of anarchy is tight: there are examples where the

quickest flow saturates every edge and the time price of anarchy is $e/(e-1)$ [Koch and Skutella 2009].

The idea of our proof is based on modifying the routing game to obtain a routing game with a single phase at equilibrium. We show that the price of anarchy in the original game is bounded by the price of anarchy in the modified game. In the equilibrium flow in the original game, we show that there exists a critical event so that in a modified equilibrium flow which has the same sequence of events, but where all the events preceding the critical event occur at time 0, and all the later events coincide with the critical event, the time price of anarchy increases. Such an equilibrium may not correspond to an actual game, as we require events which were earlier separated in time to coincide. Nevertheless, we can analyze the price of anarchy in this simpler, single-phase equilibrium and obtain an upper bound on the price of anarchy in the original game.

4. OPEN QUESTIONS

Temporal routing games go beyond the traditional assumption of static flows in routing games and offer a model for rigorous analysis that is closer to those used in traffic analyses. While the removal of the static assumption makes these games more complicated, our work shows that it is possible to obtain positive results on the price of anarchy in these games. An open question from our work is to obtain bounds on the time price of anarchy for general edge capacities, i.e., without assuming that the quickest flow saturates every edge. While our bounds on the time price of anarchy are tight, we believe that better bounds on the total delay price of anarchy are possible, and obtaining a bound on the time price of anarchy in general networks would help in refining the bound on the total delay price of anarchy.

Another interesting open problem is to compute the thin flows in each phase. In the first phase, when all the queues are increasing, the thin flow can be computed in polynomial-time [Koch and Skutella 2009]. The question of whether there exists a polynomial-time algorithm for obtaining thin flows in later phases remains open. An algorithm for computing thin flows would give us an algorithm for computing the equilibrium flow.

If we modify some of the assumptions of our model to allow for either players to control significant flow, or multiple sources, then the existence of equilibrium is an open question. Removing these assumptions would be a significant step towards enhancing the scope of applicability of routing games in the flow-over-time model.

REFERENCES

BHASKAR, U., FLEISCHER, L., AND ANSHELEVICH, E. 2011. A Stackelberg strategy for routing flow over time. In *SODA*. 192–201.

COLE, R., DODIS, Y., AND ROUGHGARDEN, T. 2006. Bottleneck links, variable demand, and the tragedy of the commons. In *SODA*. 668–677.

COMINETTI, R., CORREA, J., AND LARRE, O. 2011. Existence and uniqueness of equilibria for ows over time. In *ICALP*. To appear.

FORD, L. R. AND FULKERSON, D. R. 1962. *Flows in Networks*. Princeton University Press.

KOCH, R. AND SKUTELLA, M. 2008. Nash equilibria and the price of anarchy for flows over time. Tech. rep. Available at http://opus4.kobv.de/opus4-matheon/frontdoor/index/index/docId/530.

Koch, R. and Skutella, M. 2009. Nash equilibria and the price of anarchy for flows over time. In *SAGT*. 323–334.

Macko, M., Larson, K., and Steskal, L. 2010. Braess's paradox for flows over time. In *SAGT*. 262–275.

Mahmassani, H. S., Sbayti, H., and Zhou, X. *Dynasmart-P Version 1.0 User's Guide*. Maryland Transport Initiative, College Park, Maryland, 2004.

Roughgarden, T. 2003. The price of anarchy is independent of the network topology. *J. Comput. Syst. Sci. 67,* 2, 341–364.