European Journal of Operational Research 255 (2016) 496–504

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

A strategic timing of arrivals to a linear slowdown processor sharing system

Liron Ravner[a,*], Moshe Haviv[a], Hai L. Vu[b]

[a] Department of Statistics and the Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Israel
[b] Intelligent Transport Systems Lab, Swinburne University of Technology, Australia

ARTICLE INFO

Article history:
Received 12 August 2015
Accepted 17 May 2016
Available online 26 May 2016

Keywords:
Arrival time game
Processor sharing
Linear slowdown
Sequential timing game
Piecewise-convex programming

ABSTRACT

We consider a discrete population of users with homogeneous service demand who need to decide when to arrive to a system in which the service rate deteriorates linearly with the number of users in the system. The users have heterogeneous desired departure times from the system, and their goal is to minimise a weighted sum of the travel time and square deviation from the desired departure times. Users join the system sequentially, according to the order of their desired departure times. We model this scenario as a non-cooperative game in which each user selects his actual arrival time. We present explicit equilibria solutions for a two-user example, namely the Subgame Perfect and Cournot Nash equilibria and show that multiple equilibria may exist. We further explain why a general solution for any number of users is computationally challenging. The difficulty lies in the fact that the objective functions are piecewise-convex, i.e., non-smooth and non-convex. As a result, the minimisation of the costs relies on checking all arrival and departure order permutations, which is exponentially large with respect to the population size. Instead we propose an iterated best-response algorithm which can be efficiently studied numerically. Finally, we compare the equilibrium arrival profiles to a socially optimal solution and discuss the implications.

1. Introduction

The strategic timing of arrivals to congested systems is relevant for various applications such as traffic, queueing and communication networks. We study a non-cooperative game in which atomic users need to time their arrival to a deterministic processor sharing system with linear slowdown. This may be the case on a ring road around a business district in which the density at any point on the road affects all of the road, and therefore arriving users can cause a slowdown even for users who arrived before them. Throughout the paper we refer to the model as a traffic network where users travel along a route at varying speeds. Nevertheless, our aim here is to provide a general analysis of the strategic arrival times to such a processor sharing system, and the results are not limited to the specific traffic application. The linear slowdown dynamic can be seen as a discrete variation of Greenshield's fluid model (see for example Mahmassani & Herman, 1984). This work complements (Ravner & Nazarathy, 2015) where the socially optimal arrival schedule of users to the same system was analysed. In this paper the choice of arrival times is made by the users themselves sequentially, according to the their desired departure times. Note that while all users are served simultaneously, the model presented here still maintains the First-In-First-Out property, and thus in the sequential game users leave the system in the same order they arrived.

The study of departure time choice to a congested bottleneck goes back to Vickrey (1969), where a fluid queue dynamic was assumed. The research of fluid bottleneck models has evolved greatly since then, and we refer the reader to Arnott, de Palma, and Lindsey (1993) and de Dios Ortúzar and Willumsen (2011) and references therein. Otsubo and Rapoport analysed a non-fluid (atomic user) game with discrete arrival instances in Otsubo and Rapoport (2008). An arrival time and route choice (dynamic user equilibrium) game for a route with linear slowdown was analysed using a mean field approach by Mahmassani and Herman in Mahmassani and Herman (1984).

A queueing theory approach to the strategic timing of arrivals to a congested stochastic queue was developed by Glazer and Hassin (1983). They introduced a game in which a discrete population of users, of a Poisson distributed size, choose arrival times to a single server exponential queue with the goal of minimising waiting times. This led to another branch of research that relies on

* Corresponding author. Tel.: +972 2 5883790.
E-mail address: liron.ravner@mail.huji.ac.il (L. Ravner).

http://dx.doi.org/10.1016/j.ejor.2016.05.033
0377-2217/

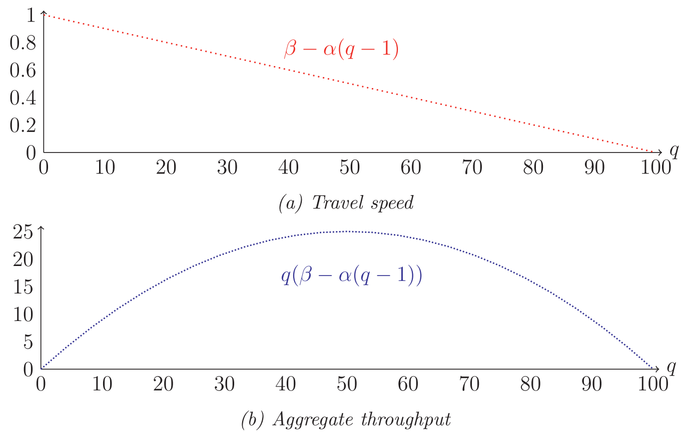

Fig. 1. The individual travel speed and aggregate throughput as the number of users in the system increases. Parameter values: $\beta = 1$ and $\alpha = 0.01$.

the stochastic properties of the queues, rather than fluid dynamics. Examples with a discrete deterministic population, as we assume in this work, are the works of Juneja and Shimkin (2013) (tardniess costs), Ravner (2014) (order penalties), and Haviv and Ravner (2015) (loss system). All of the above assume random memoryless service times and a first come first served regime. An arrival time game to a processor sharing queue was studied by Juneja and Raheja (2015), using a fluid approximation. In the queueing context our work is the first to analyse an arrival time game with a discrete population queue with heterogeneous users and deterministic service times.

2. Traffic model

In this section we introduce the model and show that despite its seeming simplicity, it in fact yields a complex arrival-departure dynamic. Suppose a set of atomic users $\mathcal{N} := \{1, \dots, N\}$ need to travel on a single route of length 1. We define the travel speed on a segment at time t as:

$$v(t) := \beta - \alpha(q(t) - 1), \tag{1}$$

where $q(t)$ is the number of users on the segment at time t, $\beta > 0$ is the free flow speed of a single user travelling alone, and $\alpha \geq 0$ is the slowdown parameter. We assume that $\beta - \alpha(N-1) > 0$, in other words, this means that the travel speed is positive even if all users travel at the same time. Note that for $\alpha \geq \frac{\beta}{2}$, the only possible case is $N = 2$. In Fig. 1 the system dynamics are illustrated as a function of the number of concurrent users. Observe that while travel speed decreases by definition, the overall service rate is non-monotone and concave. In particular, it is initially increasing, has a maximal throughput at level $q = \frac{\beta+\alpha}{2\alpha}$ and then decreases to almost zero when the system is very busy.

Every user $i \in \mathcal{N}$ has a desired departure time from the system, denoted by d_i^*. Without loss of generality we assume that the desired departure times are ordered: $d_i^* \leq d_j^*$, $\forall i < j$. The action of user i is choosing an arrival time $a_i \in \mathbb{R}$. Denote the arrival and departure vectors of all users by $\mathbf{a} := (a_1, \dots, a_N)$ and $\mathbf{d} := (d_1, \dots, d_N)$, respectively. The cost incurred by user i is

$$c_i(\mathbf{a}) = (d_i - d_i^*)^2 + \gamma\,(d_i - a_i). \tag{2}$$

This cost function is a combination of a quadratic penalty for any deviation from the desired departure time, be it early or late, and a linear penalty for the total travel time. We focus on this cost function for the sake of a clear presentation, but all of our analysis can be applied to any convex function (of the deviation and travel time terms) in a straightforward manner. We further discuss on how this generalisation can be made in the concluding remarks in Section 7. The minimal cost of any user is $\frac{\gamma}{\beta}$, and can only be obtained by travelling alone at free flow speed and leaving at exactly the desired time, d_i^* for user i.

The effective departure times of users are determined by $\mathbf{a}$ and the travel dynamics defined in (1):

$$1 = \int_{a_i}^{d_i} v(t)dt, \qquad i \in \mathcal{N}, \tag{3}$$

where $q(t) = \sum_{i \in \mathcal{N}} \mathbb{1}_{\{t \in [a_i, d_i]\}}$. Using (1) we get a set of N equations for $\mathbf{d}$,

$$1 = (d_i - a_i)(\beta + \alpha) - \alpha \int_{a_i}^{d_i} \sum_{j \in \mathcal{N}} \mathbb{1}\{t \in [a_j, d_j]\}dt, \qquad i \in \mathcal{N}.$$

These N equations can be treated as equations for the unknowns $\mathbf{d}$, given $\mathbf{a}$ or vice-versa. In Ravner and Nazarathy (2015) it was shown that the departure dynamics for an ordered vector $\mathbf{a}$ are given by

$$D\,\mathbf{d} - A\,\mathbf{a} = \mathbf{1},$$

where $A \in \mathbb{R}^N$ and $D \in \mathbb{R}^N$ are defined as follows:

$$A_{ij} := \begin{cases} \beta - \alpha(i - h_i), & i = j \\ -\alpha, & i + 1 \leq j \leq k_i \\ 0, & o.w. \end{cases}, \text{ and}$$

$$D_{ij} := \begin{cases} \beta - \alpha(k_i - i), & i = j \\ -\alpha, & h_i \leq j \leq i - 1 \\ 0, & o.w. \end{cases},$$

with $k_i := \max\{k \in \mathcal{N} \colon a_k \leq d_i\}$ and $h_i := \min\{h \in \mathcal{N} \colon d_h \geq a_i\}$. A direct result of this is the recursive formula

$$d_i = \frac{1 + (\beta - \alpha(i - h_i))a_i + \alpha\left(\sum_{j=h_i}^{i-1} d_j - \sum_{j=i+1}^{k_i} a_j\right)}{\beta - \alpha(k_i - i)}, \quad i \in \mathcal{N}. \tag{4}$$

Using an iterative algorithm we can compute the unique $\mathbf{d}$ for any given $\mathbf{a}$ (or vice versa) with at most $2N$ computations. At this point it is important to observe that the vector $\mathbf{k} := (k_1, \dots, k_N)$ defines the combined order permutation of all arrivals and departures. For example, if $N = 3$, the profile $(a_1 < a_2 < d_1 < a_3 < d_2 < d_3)$ corresponds to $\mathbf{k} := (2, 3, 3)$. In this example there is an overlay between users 1 and 2 during the interval $[a_2, d_1]$, and between users 2 and 3 during the interval $[a_3, d_2]$. Therefore, given $\mathbf{k}$ we know the exact number of users and their travel speed at any point in time. Furthermore, as long as there is no change in the permutation, the departure times are continuous with the arrival times with a known linear coefficient, as shown in Fig. 2. Note that (4) simply solves the traffic dynamics without any consideration of the cost function (2) that users wish to minimise.

If we denote the set of all possible arrival-departure permutations by

$$\mathcal{K} := \{\mathbf{k} \in \mathcal{N}^N \colon k_N = N,\ k_i \leq k_j\ \forall i \leq j\},$$

then $|\mathcal{K}| = \frac{\binom{2N}{N}}{N+1}$. This follows by observing that the elements of $\mathcal{K}$ correspond uniquely to lattice paths in the $N \times N$ grid from bottom-left to top-right with up and right movements without crossing the diagonal. The number of such elements is the Catalan number (see Ravner and Nazarathy, 2015 for more details).

The relation between arrival and departure times, defined in (3), is in fact a set of piecewise-linear equations. This is illustrated in Fig. 2 by changing the arrival time of a single user while keeping all others fixed.

The piecewise linear relation between arrival and departure times implies that the cost of any single user is not convex with respect to his own arrival time, even though the cost function, defined in (2), has a convex form. This piecewise-convex behaviour is illustrated for the same numerical example in Fig. 3. We can see