Dynamic Service Sharing with Heterogeneous Preferences

Israel Ben-Shahar, Ariel Orda and Nahum Shimkin

Department of Electrical Engineering
Technion, Israel Institute of Technology, Haifa 32000, Israel
E-mail: {ibs@tx;ariel@ee;shimkin@ee}.technion.ac.il

We consider a service system where individual users share a common resource, modeled as a processor-sharing queue. Arriving users observe the current load in the system, and should decide whether to join it or not. The motivation for this model is based, in part, on best-effort service classes in computer communication networks. This decision problem is modeled as a noncooperative dynamic game between the users, where each user will enter the system only if its expected service time (given the system description and policies of subsequent users) is not larger than its quality of service (QoS) requirement. The present work generalizes previous results by Altman and Shimkin (1998), where all users were assumed identical in terms of their QoS requirements and decision policies; here we allow heterogeneous requirements, hence different policies. The main result is the existence and uniqueness of the equilibrium point in this system, which specifies a unique threshold policy for each user type. Computation of the equilibrium thresholds are briefly discussed, as well as dynamic learning schemes which motivate the Nash equilibrium solution for this system.
Keywords: Resource Sharing, Processor Sharing Queues, Individual Optimality, Nash Equilibrium, Dynamic Games, Best-Effort Service.

1. Introduction

We consider a class of service systems where all users are served in parallel and in equal shares, in accordance with the Processor-Sharing service discipline. Due to this sharing of service resources, the service rate to each individual user decreases as the load (number of concurrent users) on the system increases. The aggregate service quality perceived by an individual user thus depends on the system load profile throughout its service period.

Each arriving user is assumed to observe the system load at the arrival instant. Based on that information, the user needs to decide whether to join or leave the system; in the latter case it may abandon the required service altogether, or perhaps seek an alternative service option. A natural performance measure on which the decision whether to join or leave may be based on is the expected service time (which coincides with the sojourn time in this system): a user should form an estimate of its expected service time, given the perceived system load, and join the system only if the estimated time is satisfactory, i.e., smaller than some user-dependent threshold.

The situation just described is relevant, for example, to *Best Effort* service classes in communication networks, including the current Internet, where the available bandwidth is shared by all users. Further discussion of this application may be found in [5], and several others, ranging from sharing a central computer system to transportation and public recreational activities, are mentioned in [2]. Depending on the application, the observation of the system state by an arriving user may be performed via an actual observation of other users, a query of the system log, or simply by perceiving its own initial service rate as an indication for the current system load.

A simple but naive approach for estimating the expected service time would be based on the assumption that the initial load persists throughout a user's service period. This assumption ignores of course the stochastic variation in the system state, which will be significant especially if the user spends a relatively long time in the system. The naive approach will also be at odds with acquired experience, as suggested by the learning scenario discussed in Section 7.

More realistically, as the service duration for each user is affected by the system load throughout its service period, the effect of subsequent arrivals must be taken into account. But these arrivals are themselves individual users which are part of the overall decision problem, hence their decisions will affect the performance of the users already in service. This dependence leads us to consider this individual decision problem as a noncooperative game [14] among users, and study the Nash equilibrium solution of this game. We shall consider the existence, uniqueness, structure and computation of the equilibrium point. Of particular interest here will be the uniqueness issue; this is not inherent in the Nash equilibrium solution, and in general multiple unconnected equilibria may exist. We shall also consider some learning schemes which motivate the Nash equilibrium solution, which emerges as the stationary operating point of the resulting dynamical system under a reasonable learning scenario.

A similar model to the one studied here has been considered in [2], where it was assumed that users are homogeneous in all respects, including their expected service time requirements, hence in their equilibrium policies. In the present study we remove this complete homogeneity assumption, and consider the case of heterogeneous users which differ in their expected service time requirements. Consistent with standard terminology in modern communication networks, we shall refer to the latter as the QoS (Quality of Service) requirement.

In the homogeneous case considered in [2], the uniqueness question is naturally reducible to a scalar problem. Any possible equilibrium policy profile may be characterized by a single threshold value (common to all users), and represented by a real number. The set of all equilibrium points is thus parameterized by a single (real-valued) threshold, and its uniqueness may be deduced from basic monotonicity properties of the expected waiting time. On the other hand, in the heterogeneous case considered here, different QoS requirements give rise to different policies for different users, hence the problem becomes multi-dimensional in nature. To establish uniqueness, we resort here to a more elaborate comparison argument, which relies on stochastic coupling in order to compare directly the expected waiting times under different policy profiles. We note that the present uniqueness proof is valid only for Poisson arrivals, while [2] treats GI arrivals.

Some related work on game-theoretic analysis of Processor Sharing queues may be found in [3,9]. The first considers the problem of reneging from the queue by homogeneous customers, and the second concerns the problem of joining the queue in the absence of state information.

Game theoretic models for user behavior in queueing and networking problems have been the subject of considerable research during the past few years, e.g. in the context of flow control [10]–[11], routing [8,16], virtual path bandwidth allocation [12], and queue abandonments models [13]. These studies provide valuable insight into the qualitative and quantitative aspects of system operation in a user-driven environment.

Our model is also closely related to problems of admission control to a single queue [15]–[17]. Under various cost and queueing conditions it has been established that the socially optimal admission policy is a threshold policy (with respect to the queue length at arrival). These social policies have often been compared to the individually optimal ones under a *FCFS* service discipline. Under this service discipline the determination of the individually optimal policy for

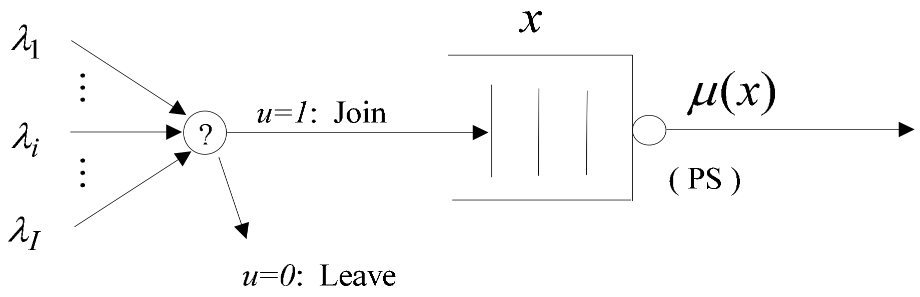

Figure 1. The processor-sharing queueing model. Users of type i arrive at a rate $\lambda_i = p_i\lambda$, and are characterized by a QoS requirement parameter θ^i.

each user becomes trivial, since, given the queue state information, the expected sojourn time does not depend on actions of other users. This dependence does exist under any other service discipline, including the processor sharing discipline considered here.

The paper is organized as follows. The basic model in formulated in Section 2. Existence of a Nash equilibrium and some structural properties thereof are derived in Section 3. In particular, it is shown that any best-response policy (hence any equilibrium policy) must be a threshold policy. Uniqueness of the equilibrium point is established in Section 4. Section 5 provides an iterative scheme for computing the equilibrium point. The generalization from a finite to an infinite (continuous) set of QoS parameters is the subject of Section 6. Dynamic learning schemes are presented and discussed in Section 7. The paper closes with some concluding remarks in Section 8.

2. The Model

Consider a processor-sharing service system, which equally divides its service resources among all current users. The user arrival process is Poisson, with rate λ. Each arriving user may choose to join the system or leave immediately, based on criteria to be specified below; see Figure 1. The system has a buffer of size B, which may be finite or infinite.

The service requirement of each user is an exponential random variable W with mean 1, independent across users, and measured in units of work (or “bits”). The service rate obeys the following assumptions:

(a) The server rate is $\mu(x)$ (bits/sec), where $x \geq 1$ is the number of users in the system.
(b) $0 < \mu(x) \leq \mu_{max}$ for a finite bound μ_{max}.
(c) The individual service rate applied to each user is $\nu(x) := \mu(x)/x$.
(d) The individual service rate $\nu(x)$ is strictly decreasing in $x \geq 1$.

The last assumption means that the service rate to each user decreases as the load increases. This coupling between users is at the core of our resource sharing problem. Note that the total service rate $\mu(x)$ is also allowed to depend on the system load, as may be the case in many systems of interest. The form of this dependence is not important, as long as $\nu(x)$ is decreasing.

Let $X(t)$ denote the number of users in the system at time t. A user that joins the system at time t_o will stay until its entire service requirement W is fulfilled, namely the first time t such that

$$\int_{t_o}^{t} \nu(X(s))\, ds = W \ .$$

Users are divided into a finite set $\mathcal{I} = \{1, \ldots, I\}$ of *types*, according to their quality of service (QoS) requirements. A user of type i is characterized by a parameter θ^i, which denotes its QoS requirement, i.e., the required service time for this type. We assume that $\theta^i \neq \theta^j$ for $i \neq j$. Essentially, an arriving user of type i will find the offered service acceptable if $E(V) \leq \theta^i$, where V is its total service time and E an expectation conditioned on available information and system state; we shall be more precise below. The type $i \in \mathcal{I}$ of each user is determined randomly, and independently of all other user characteristics. Let p_i denote the probability that an arriving user is of type i, with $\sum_{i \in \mathcal{I}} p_i = 1$.

It will be convenient to let the system start at time $t = 0$, with some initial state $X(0) = x_o$. Let T_k, $k \geq 0$, denote the arrival time of the k-th user, with $0 = T_o < T_1 < T_2 \cdots$. $X(T_k)$ is then the queue size just prior to the arrival of user k. Denote by i_k the type of user k.

We next describe the individual decision structure. Upon its arrival (at T_k), user k may decide either to join the system, or leave it without returning. It is assumed that the user observes the queue length $X(T_k)$ at its arrival, and may base its decision on that information. A *decision rule*, or policy, for user k is a function $u_k : \{0, 1, \ldots, B-1\} \to [0, 1]$. This assigns to each possible queue size $X(T_k) = x$ a randomized decision $u_k(x) \in [0, 1]$, which specifies the probability of joining the system at that queue size. Note that the buffer is full when $x = B$, so that a user that arrives at that queue size cannot join the queue and must leave. A *policy profile* $\pi = (u_o, u_1, \ldots)$ is a collection of decision rules for all arriving users. We may now define $V_k(x, \pi)$ – the expected service time for user k, given that x other users are present upon its arrival, and that all other (subsequent) users follow the policy profile π.

A policy u_k is defined to be a *best response* of user k against π if

$$u_k(x) = \begin{cases} 1 & : V_k(x,\pi) < \theta^{i_k} \\ 0 & : V_k(x,\pi) > \theta^{i_k} \\ \text{any } q \in [0,1] & : V_k(x,\pi) = \theta^{i_k} \end{cases} \tag{1}$$

We also refer to such a policy as *individually optimal.* Thus, a user chooses to enter the system if the expected service time is not larger than the required service time, which is given by the parameter θ^i that characterizes its type i. In case of a tie, the user is neutral between joining the system or leaving, so that it may join with any probability $0 \leq q \leq 1$.

We note that this definition of individual optimality corresponds directly to the interpretation of θ^i as a QoS requirement on the expected service time. However, this definition may be equivalently interpreted in a more standard way as the cost-minimizing choice with respect to an appropriate cost function, in which the cost associated with joining the system is just the expected service time there, while the cost for leaving is θ^i (cf. [2]). Moreover, if we interpret θ^i as an equivalent service time associated with some alternative to the processor sharing system, then this cost function is just the expected service time of the user given its decision.

A policy profile $\pi = (u_o, u_1 \ldots)$ is a *Nash equilibrium point* (or just an *equilibrium point*) for the system if each policy u_k is a best response policy for user k against π. These equilibria will be the main object of study in this paper.

We note that in game theoretic terms the equilibrium just defined is a *subgame-perfect* equilibrium point, since the decisions of users are specified in all possible instances of their information sets (i.e., queue lengths). This is consistent with the standard definitions in stochastic games. We further note that we require each policy to depend on the current system state but not on its history. Arguably, this is the natural information structure in the applications considered here.

A policy profile is said to be *type-homogeneous*, or just *homogeneous*, if all users of the same type employ the same policy. An equilibrium point is homogeneous if it consists of a homogeneous policy profile. We shall focus in this paper on homogeneous policies and equilibria, which reflect the *a-priori* assumption that users with identical parameters and preferences will behave identically. In fact this assumption can be considerably relaxed, see the remark at the end of Section 4.

We further define a *threshold policy* as follows: given $0 \leq q \leq 1$ and $L \geq 0$, an (L, q)-threshold policy u is defined as:

$$u(x) = \begin{cases} 1 & \text{if } x < L \\ q & \text{if } x = L \\ 0 & \text{if } x > L \end{cases}$$

An (L, q)-threshold policy will be denoted as $[L, q]$, or simply $[g]$ with $g = L + q$. When the buffer size is finite, the policy $[g]$ with $g > B$ is equivalent to $[B]$.

A *homogeneous threshold profile* is then specified by the I-tuple of thresholds $g = (g^1, \ldots, g^I)$, where g^i is the threshold policy to be employed by users of type i. We shall identify the vector g with the policy profile implied by it; thus, we may consider g either as a vector in $\mathbb{R}_+^I$ or as a policy profile, depending on the context.

3. Existence and Basic Structure

In this section we derive a basic structural property of an equilibrium point, namely that any of its component policies must be a threshold policy. This follows directly from the monotonicity of the expected service time $V_k(x, \pi)$ in x. We shall then establish the existence of a homogeneous equilibrium point, using some continuity properties of the best response map. The required monotonicity and continuity properties of V_k are summarized in Lemmas 1 and 3. These properties are direct extensions of the corresponding results in [2] for the single-type case; their proofs, which rely on basic stochastic coupling arguments, are omitted here.

Lemma 1 (see [2]). $V_k(x, \pi)$ is strictly increasing in $x \geq 0$, for every policy profile π and user k.

As a direct consequence, we have the following property of any best-response policy.

Proposition 2 (best-response policies).

(i) Any best-response policy (for user k against policy profile π) is a threshold policy.
(ii) For user k of type i, the set $U^i_k(\pi)$ of best-response policies against π is given as follows. Let $L^i = \min\{L \in \mathbb{Z}_+ : V_k(L^i, \pi) \geq \theta^i\}$. Now,
(a) if $V_k(L^i, \pi) = \theta^i$, then $U^i_k(\pi) = \{[L^i, q^i] : \ 0 \leq q^i \leq 1\}$,
(b) otherwise $U^i_k = \{[L^i, 0]\}$, i.e. a singleton.

As in [2], the proof is an immediate consequence of (1), combined with the monotonicity property in the previous lemma. We may now conclude, by the definition of equilibrium, that any homogeneous equilibrium point is a homogeneous threshold profile specified by a threshold vector $g = (g^1, \ldots, g^I)$.

The next lemma concerns the continuity and monotonicity of the service times $V_k(x, g)$ in g, which will be required in the existence proof to verify the continuity of the best response map.

Lemma 3. For every $k \geq 0$, $x \geq 0$ and $g \in [0, B)^I$, $V_k(x, g)$ is

(i) strictly increasing in each component of g,
(ii) continuous in g.

The proof is similar to that of Lemma 2 in [2], applied to each component of g.

Theorem 4 (existence). For the model described in Section 2, there exists a homogeneous equilibrium point, which consists of a homogeneous threshold profile g.

Proof. Let $\widetilde{\Gamma}$ denote the set of threshold policies, $\widetilde{\Gamma} = \{g : g = (g^1, \ldots, g^I) : g^i \in [0, B)\}$, which is a subset of $\mathbb{R}^I$. Let G be the homogeneous best-response map over $\widetilde{\Gamma}$; according to Proposition 2, $G(g) = \times_{i=1}^I U^i_k(g) \subset \widetilde{\Gamma}$. Note that $G(g)$ is a point-to-set mapping. It is required to show that G has a fixed point g_o, namely $g_o \in G(g_o)$, which is then obviously an equilibrium point of the required form. For that purpose we employ the following version of the Kakutani fixed-point theorem [4]: Let S be a compact and convex subset of $\mathbb{R}^n$, and let f be an upper semi-continuous point-to-set map over S, so that $f(x) \subset S$ is closed for every x. Then $x \in f(x)$ for some $x \in S$.

Since $\widetilde{\Gamma}$ may be unbounded in case the threshold B is infinite, we shall consider an appropriate bounded subset. Observe that the threshold value of any best-response policy for type i ($i \geq 1$) is uniformly bounded. Indeed, since the service time increases to infinity as the queue size increases, there exists a threshold g^i_{max} for each type i, beyond which it is never optimal to enter. Consider then the restriction of $\widetilde{\Gamma}$ to these values, denoted Γ, which is a compact convex subset of $I\!R^I$, and satisfies $G(g) \subset \Gamma$ for every g. Obviously $G(g)$ is closed, and upper semi-continuity of G follows easily from (1) and Lemma 3(ii). Thus there exists a fixed point $g_o \in G(g_o)$, which is a homogeneous threshold equilibrium by its definition. □

4. Uniqueness

The previous section established that any homogeneous equilibrium point must be a threshold profile, and that at least one such equilibrium point exists. Here we establish the uniqueness of that equilibrium.

Theorem 5 (uniqueness). The homogeneous equilibrium point is unique.

The proof of this theorem proceeds through several lemmas. The first two concern some useful monotonicity properties of the equilibrium thresholds.

Throughout this section, we shall be concerned with homogeneous threshold profiles $g = (g^1, \ldots, g^I)$. It is easily seen that under such policy profiles the expected service time $V_k(x, g)$ does not depend on the user index k, which will thus be omitted.

Lemma 6. Let $g = (g^1, \ldots, g^I)$ be a homogeneous equilibrium point. Then $\theta^i > \theta^j$ implies $g^i \geq g^j$.

Proof. By definition of an equilibrium point, g^i and g^j are best responses against g for user types i and j, respectively. The assertion follows from (1) and the monotonicity of $V(x, g)$ in x, from in Lemma 1. □

Lemma 7. Let g_a and g_b be two homogeneous threshold equilibrium profiles. Then $g_a \geq g_b$ implies $g_a = g_b$.

Proof. Assume, by way of contradiction, that $g_a \geq g_b$, and $g^i_a > g^i_b$, for some i. By Lemma 3(i) it follows that $V(x, g_a) > V(x, g_b)$, and Proposition 2(ii) then implies that $g^i_a \leq g^i_b$, a contradiction. □

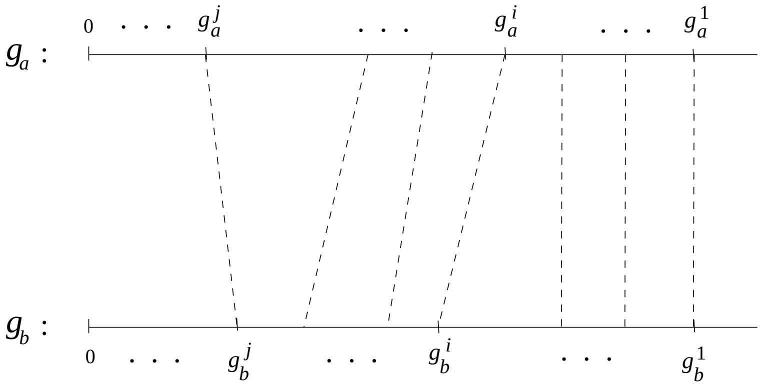

Figure 2. The relation between the two threshold profiles in the proof of Theorem 5: $g_a^k = g_b^k$ for $1 \leq k < i$, $g_a^i > g_b^i$, $g_a^k \geq g_b^k$ for $i < k < j$, and $g_a^j < g_b^j$.

To establish Theorem 5, we shall assume existence of two different equilibrium points, and show that this leads to a contradiction. Recall, from Proposition 2, that any homogeneous equilibrium point must be a threshold profile g. We thus need to consider here only policy profiles of that form. To simplify notation, observe that, under a homogeneous threshold profile g, the expected service time $V_k(x, g)$ does not depend on the user index, which we drop below.

Assume now that $g_a = (g_a^1, \ldots, g_a^I)$ and $g_b = (g_b^1, \ldots, g_b^I)$ are two equilibrium points, and $g_a \neq g_b$. By Lemma 7, neither one can completely dominate the other, hence there exist some type indices i and j so that $g_a^i > g_b^i$ and $g_a^j < g_b^j$. Let us re-order the user types so that $\theta^1 > \theta^2 > \cdots > \theta^I$. Let now i be the *smallest* index for which $g_a^i > g_b^i$, and j be the smallest index for which $g_a^j < g_b^j$. Assume, without loss of generality, that $i < j$, that is $\theta^i > \theta^j$. Lemma 6 then implies that $g_b^j \leq g_b^i$. These relations between the thresholds are illustrated in Figure 2. Note, in particular, that $g_a^j < g_b^j \leq g_b^i < g_a^i$.

Let $L_b^i = \lfloor g_b^i \rfloor$ denote the integer part of g_b^i. We shall establish the following contradiction. On the one hand, using the inequality $g_a^i > g_b^i$ and the respective optimality of these thresholds, it will be argued that $V(L_b^i, g_a) \leq V(L_b^i, g_b)$. On the other hand, we shall argue that $V(L_b^i, g_a) > V(L_b^i, g_b)$. This will be performed by comparing the two stochastic processes that evolve under the policy profiles g_a and g_b. Obviously, these inequalities will contradict the assumption that $g_a \neq g_b$, and imply the uniqueness of the equilibrium. The first inequality is established in the following lemma.

Lemma 8. $g_a^i > g_b^i$ implies $V(L_b^i, g_a) \leq V(L_b^i, g_b)$, where L_b^i is the integer part of g_b^i.

Proof. From Proposition 2(ii) it follows that L_b^i is the smallest integer such that $V(L_b^i, g_b) \geq \theta^i$ holds, and L_a^i is the smallest integer for which $V(L_a^i, g_a) \geq \theta^i$. Now $g_a^i > g_b^i$ implies that $L_a^i \geq L_b^i$. If $L_a^i > L_b^i$ it follows by Lemma 1 that $V(L_b^i, g_a) < \theta^i$, and consequently $V(L_b^i, g_a) < V(L_b^i, g_b)$. If $L_a^i = L_b^i$, then $g_a^i > g_b^i$ implies that g_a^i is non-integer, and, by Proposition 2(ii), $V(L_b^i, g_a) = V(L_a^i, g_a) = \theta^i$. Thus, using again $V(L_b^i, g_b) \geq \theta^i$, it follows that $V(L_b^i, g_a) \leq V(L_b^i, g_b)$. □

We next turn to the more involved task of showing that $V(L_b^i, g_a) > V(L_b^i, g_b)$ holds under our assumptions. The idea is as follows. Consider a user that joins the queue at queue-length L_b^i, under each threshold profile. Stochastic coupling will be used to compare the stochastic processes that evolve in both cases, to which we refer as case (a) and case (b), respectively. The service time of this user will be considered separately over two periods.

(i) Up until the first time r^j that the queue length in either case (which turns out to be always case (b)) first hits $\hat{L} := L_b^j + 1$: since $g_a^\ell \geq g_b^\ell$ for $\ell < j$, more users may enter the queue in case (a) than in case (b), the queue tends to be more congested in case (a), and therefore the considered user tends to stay longer in case (a).

(ii) From time r^j onward: Here we start with the relation $g_a^j < g_b^j$ which, using the optimality of these thresholds, implies that $V(L_b^j, g_a) > V(L_b^j, g_b)$. Observe further that the queue length at r^j equals $L_b^j + 1$ in case (b), and is at least as large in case (a). Thus, provided that our user is still in the queue at time r^j, the last inequality can be used to compare its *remaining* service time beyond r^j in both cases.

We proceed to make this argument more rigorous. We start by defining the two processes of interest and their stochastic coupling.

Let $X_a(t)$ denote the queue-length process under the threshold profile g_a, with initial conditions $X_a(0^-) = L_b^i$, and assuming that user C_0 joins the queue at time 0. Note that the expected service time for this user is just $V(L_b^i, g_a)$. Similarly, let $X_b(t)$ denote the queue-length process under the threshold profile g_b, with the same initial conditions and user C_0, which now has $V(L_b^i, g_b)$ as its expected service time. We refer to the two processes just defined as cases (a) and (b), respectively.

We now use stochastic coupling to define the two processes on a common probability space, which allows their sample-path comparison. The coupled processes will be defined so that the following hold.

(P1) External arrival times of users of each type are identical in both cases.

(P2) For each arriving user, a higher probability of joining is translated to sample-path dominance; that is, if a certain user decides (according to its policy and the queue size) to join with probability q_a in case (a), and probability q_b in case (b), then $q_a > q_b$ (say) implies that this user must join in case (a) if it does in case (b).

(P3) On times t when $X_a(t) = X_b(t)$, the service completion times in cases (a) and (b) coincide.

(P4) At any time t when $\nu(X_a(t)) \leq \nu(X_b(t))$ i.e. $X_a(t) \geq X_b(t)$, a service completion of user C_0 in case (a) at t implies a service completion of C_0 in case (b) at or before t.

We note that property (P4) may be implemented due to the memoryless nature of the service process. A detailed construction of the probability space and stochastic processes that satisfy the stated properties may be found in [2].

Let $\hat{L} = L_b^j + 1$, with $L_b^j = \lfloor g_b^j \rfloor$ the integer part of g_b^j, and define $r^j = \inf\{t > 0 : X_b(t) = \hat{L}\}$. Note that $r^j = 0$ if $L_b^i = L_b^j$ (since $X_b(0^+) = L_b^i + 1$), and $r^j > 0$ otherwise. Further denote the service time of user C_0 in cases (a) and (b) as r_a and r_b, respectively. We divide these times into two components, as follows:

$$V(L_b^i, g_a) = E(r_a) = E(r_a \mathbf{1}\{r_b \leq r^j\}) + E(r_a \mathbf{1}\{r_b > r^j\}) \tag{2}$$

$$V(L_b^i, g_b) = E(r_b) = E(r_b \mathbf{1}\{r_b \leq r^j\}) + E(r_b \mathbf{1}\{r_b > r^j\}) \tag{3}$$

Here $\mathbf{1}\{\cdot\}$ is the indicator function.

We shall compare each of these components separately. Lemmas 9 and 10 compare the first component, based on the fact that, up to r^j, the system is more congested in case (a) than in case (b), due to the thresholds employed. Lemmas 11 and 12 treat the second component, based on the properties of the thresholds $g_a^j < g_b^j$. These results are then combined to obtain the required inequality in Lemma 13.

Lemma 9. $X_a(t) \geq X_b(t)$ for $t \leq r^j$.

Proof. Assume that $X_a(t) \geq X_b(t)$ for some $t < r^j$. We will show that this inequality persists through any single change in the queue size, namely service completion or external arrival. Noting that $X_a(0^+) = X_b(0^+) = L_b^i + 1$, this clearly implies the assertion. Let t be an event time. Noting that the queue length changes by one unit on any single event, it is obvious that $X_a(t^-) > X_b(t^-)$ implies $X_a(t^+) \geq X_b(t^+)$ after any single event. It remains thus to consider the case when $X_a(t^-) = X_b(t^-)$, which is assumed next.

In the event of a service completion, property (P3) of the stochastic coupling scheme ensures that the service completion times coincide in both cases, and $X_a(t^+) = X_b(t^+)$ persists. Consider next an external arrival, and let ℓ denote the type of the arriving user. Here we have two options:

1. If $\ell < j$, then $g_a^\ell \geq g_b^\ell$ by definition of j, the probability with which the user enters the queue in case (a) is not smaller than in case (b), and by property (P2) it follows that $X_a(t^+) \geq X_b(t^+)$.
2. If $\ell \geq j$, then the thresholds g_a^ℓ and g_b^ℓ are smaller than $L_b^j + 1$ (cf. Figure 2), while $X_a(t^-) = X_b(t^-) \geq L_b^j + 1$ by definition of r^j. Thus the user will choose not to join the queue in either case.

It follows that $X_a(t) \geq X_b(t)$ holds for $t \leq r^j$. $\square$

Lemma 10. $E(r_a \mathbf{1}\{r_b \leq r^j\}) \geq E(r_b \mathbf{1}\{r_b \leq r^j\})$ with strict inequality if $L_b^i > L_b^j$.

Proof. If $r_b \leq r^j$, then for $0 \leq t \leq r_b$ we obviously have $t \leq r^j$, hence from Lemma 9, $X_a(t) \geq X_b(t)$. But then Property (P4) of the stochastic coupling implies $r_a \geq r_b$, hence the corresponding inequality of the expected values. Furthermore, if $L_b^i > L_b^j$, we claim that $r_a > r_b$ with positive probability on $\{r_b \leq r^j\}$. Indeed, note first that $L_b^i > L_b^j$ implies $r^j > 0$ (since $X_b(0^+) \triangleq L_b^i + 1 > L_b^j + 1$), and consequently $P(r_b \leq r^j) > 0$. The following chain of events may then be seen to have a positive probability of occurrence: $X_b(t)$ reaches L_b^i at some $0 \leq t \leq r_b$, and then $g_a^i > g_b^i$ allows that some i-type user will join the queue in case (a) but not in case (b). Subsequently $X_a(t) > X_b(t)$ results in the service intensity (to user C_0 in particular) becoming higher in case (b), and $r_a > r_b$ is obtained with positive probability. Combined with $r_a \geq r_b$, this implies the required strict inequality. $\square$

The next two lemmas compare the last terms in (2) and (3).

Lemma 11. $g_a^j < g_b^j$ implies $V(L_b^j, g_a) \geq V(L_b^j, g_b)$, with strict inequality if $L_a^j < L_b^j$.

Proof. Identical to the proof of Lemma 8, where the condition for strict inequality may also be found. □

Lemma 12. $E(r_a \mathbf{1}\{r_b > r^j\}) \geq E(r_b \mathbf{1}\{r_b > r^j\})$ with strict inequality if $L_a^j < L_b^j$.

Proof. Consider case (b) first. It is claimed that

$$E(r_b - r^j | r_b > r^j) = V(L_b^j, g_b) \; . \tag{4}$$

Indeed, $r_b > r^j$ means that user C_0 is still in queue just after r^j, and by definition $X_b(r^j) = L_b^j + 1$. By the memoryless nature of the service and arrival process, the distribution of the remaining service time $(r_b - r^j)$ for that user will be identical to the distribution of its service time in case it joins the queue at time 0, at queue size $X(0^-) = L_b^j$. But the expected value of the latter is $V(L_b^j, g_b)$ by definition.

Consider case (a) next. It was shown in Lemma 10 that $r_a \geq r_b$ on $\{r_b \leq r^j\}$, so that $r_b > r^j$ implies $r_a > r^j$. Using the same consideration as above, it follows that

$$E(r_a - r^j | r_b > r^j) = E[V(X_a(r^j+) - 1, g_a) | r_b > r^j] \tag{5}$$

But by Lemma 9, $X_a(r^j+) \geq X_b(r^j+) = L_b^j + 1$, so that, by monotonicity of V in x,

$$E(r_a - r^j | r_b > r^j) \geq V(L_b^j, g_a) \; . \tag{6}$$

Observing Lemma 11, (4) and (6) now imply that

$$E(r_a - r^j | r_b > r^j) \geq E(r_b - r^j | r_b > r^j)$$

which clearly implies the required inequality.

Strict inequality under $L_a^j < L_b^j$ follows similarly by noting the strict inequality in Lemma 11, and the easily verified fact that $P(r_b > r^j) > 0$. □

Remark. The last lemma, in particular the equalities (4) and (5), rely on the memoryless nature of the arrival process. Thus, in contrast to [2] which considers i.i.d. arrivals in general, the present uniqueness result is derived strictly under Poisson arrivals.

We are now ready to prove:

Lemma 13. $V(L_b^i, g_a) > V(L_b^i, g_b)$.

Proof. Using Lemmas 10 and 12 in (2) and (3) clearly implies that $V(L_b^i, g_a) \geq V(L_b^i, g_b)$, with strict inequality unless $L_a^j = L_b^j = L_b^i$. However, recalling the starting-point relations $g_a^j < g_b^j \leq g_b^i$, this implies that g_b^j and g_b^i are non-integer. By Proposition 2(ii) and the optimality of g_a^j against g_a we then obtain $V(L_a^j, g_a) = \theta^j$, and similarly $V(L_b^i, g_b) = \theta^j$. But $\theta^i > \theta^j$ by assumption, and strict inequality follows in that case as well. $\square$

Comparing Lemma 8 and Lemma 13, it is obvious that the assumption $g_a \neq g_b$ leads to contradiction. Hence, the homogeneous equilibrium point is unique, and Theorem 5 is established. $\square$

Remark (uniqueness of non-homogeneous equilibria). Consider the following relaxation of the homogeneity assumption. Suppose that user types are further split into "sub-types", such that the homogeneity assumption is imposed only at the sub-type level. Thus, users with the same service time requirement may employ different policies. An arriving user's sub-type is assigned according to some time-independent distribution. It is then easy to see that strict equilibrium uniqueness as derived above ceases to hold. As an example, suppose that, in the original setting, users of a certain type i employ an equilibrium policy $[L, 0.5]$; recall that this implies that, at equilibrium, each user of this type is indifferent between $[L, 0]$ and $[L, 1]$, and arriving users are probabilistically split between these two options. Suppose now that these users are equally split into two sub-types i_1 and i_2, which employ policies $[L, q]$ and $[L, 1-q]$, respectively, for some $0 \leq q \leq 1$. Then, it is easy to verify that the original policy profile thus modified is also an an equilibrium policy (for any q). Nonetheless, we observe that, at a "macroscopic" level, i.e., for an observer that does not distinguish among users at the sub-type level, the type behavior is statistically the same, and the system exhibits an identical behavior at all these equilibria. In fact, this "macroscopic uniqueness" holds in general, as can be immediately deduced from our previous uniqueness results. A formal statement and proof are however outside the scope of this paper.

5. Equilibrium Computation

We next address briefly the issue of computation of the equilibrium point. By our previous results, the homogeneous equilibrium point is characterized by the set of thresholds $g = (g^1, \dots, g^I)$, which we wish to compute numerically.

We propose an iterative algorithm, which performs repeated single-type computations based on the following concept. Given a threshold profile g, define the best-response equilibrium for user-type i against g as the homogeneous equilibrium point for this type, given that all other users are known to follow the thresholds $g^{-i} = (g^1, \dots, g^{i-1}, g^{i+1}, \dots, g^I)$. This is simply characterized by a scalar threshold (see below), say g^i_*. Note that the best-response equilibrium is different from the best-response policy for a single user, as specified in (1), since here we require the threshold of all type-i users to change simultaneously.

The best-response equilibrium may be viewed as a single-type specialization of the multi-type model; it is thus easy to verify that the best-response equilibrium is a unique threshold policy, and its value (g^i_*) is continuous in g^i. Its computation is similar to the computation of the equilibrium threshold for the single user-type problem, treated in [2], and will not be repeated here.

The iterative algorithm proposed to compute the equilibrium is then:

(i) Start with some set of thresholds $g_0 = (g^1_0, \dots, g^I_0)$.

(ii) At step $n = 1, 2, \dots$ compute the best-response threshold $g^{i_n}_n$ for user-type i_n against g_{n-1} and replace to form g_n, with $\{i_n\}$ varying cyclically over the set of user types.

It is easily verified that if this algorithm converges, the limit g_∞ must be an equilibrium point. Furthermore, for the case of two user-types we have the following convergence result:

Proposition 14. Suppose $I = 2$. Then g_n converges monotonically to the equilibrium profile.

Proof. Using the monotonicity and continuity properties in Lemma 3, it may be verified that the best-response equilibrium g^1 against g^2 is non-increasing and continuous in g^2, and vice-versa. It follows that both components of g_n are monotonic (one non-increasing and the other non-decreasing), hence converge to a limit point g_∞. By continuity it then follows that each component of g_∞ is a best-response to g_∞, hence the latter is an equilibrium profile. $\square$

For $I > 2$, the algorithm invariably converged in numerous experiments. However, a proof of convergence is not available as yet.

6. The Continuous-Parameter Case

So far it has been assumed that users are divided into a finite number of types, based on a discrete set of QoS requirement parameters θ^i, $i \in \mathcal{I}$. In some cases it may be more natural to assume that θ is a continuous (non-discrete) variable, which represents the statistical variability between the preferences of different users. We briefly, and somewhat informally, consider this case in this section.

Assume then that the parameter $\theta \in \Theta \subset I\!R$ is determined for each user according to a probability measure P_Θ. Assume further that P_Θ has a finite density f_Θ on $I\!R$. (We thus avoid point masses in the distribution of Θ, and in particular preclude the previous case of a discrete parameter. A unified treatment of both cases may be given, but tends to complicate the notation and seems to be of little added value.)

A homogeneous policy profile is now a collection of policies $\{\pi_\theta, \quad \theta \in \Theta\}$. A *threshold* profile is the map $g : \Theta \to [0, \infty)$, which associates a threshold $g(\theta)$ with each $\theta \in \Theta$.

Since Lemma 1 holds, it follows that any homogeneous equilibrium profile is a threshold profile. Moreover, it may be verified that Lemma 3(i) holds as well; that is, $V(x, g)$ is strictly increasing in g, in the sense that $g_1 \geq g_2$ and $P_\Theta\{g_1(\theta) > g_2(\theta)\} > 0$ imply $V(x, g_1) > V(x, g_2)$ for every x.

Assume that g is an equilibrium point, and let $\{V(x, g) : x = 0, 1, \ldots, B-1\}$ be the corresponding service-time vector. Then, by definition of the best-response policies and since $V(x, g)$ is increasing in x, it follows that g has the following form:

$$g(\theta) = \begin{cases} L & : V(L-1, g) < \theta < V(L, g) \\ L + q, \text{ any } q \in [0, 1] & : \theta = V(L, g) \end{cases}$$

for $\theta \in \Theta$ and $L = 0, 1, 2, \ldots, B-1$.

It follows that $g(\theta)$ has a "staircase" form, starting from $g(0) = 0$ and increasing by jumps of 1 at each point $\theta_L = V(L, g)$. It follows that $g(\theta)$ can be identified with the discrete set $\{\theta_L, \ L = 0, 1, \ldots, B-1\}$. It is interesting to note that the equilibrium thresholds here are essentially deterministic (except possibly on the set $\{\theta_1, \ldots, \theta_{B-1}\}$, which has zero measure). This is in contrast to the discrete parameter case of the previous sections, where in general randomized thresholds are essential for the existence of an equilibrium point. Obviously, in the latter case randomization is required to generate a continuum of policies, while in the present (continuous parameter) case this already exists in deterministic policies.

Given the discrete representation $\{\theta_L\}$ of an equilibrium profile, existence and uniqueness of the equilibrium point may be established in the present case as well. The proofs are a straightforward adaptation of the discrete parameters case, and will not be elaborated here.

7. Learning

The Nash equilibrium solution is the "rational" choice for a sophisticated decision maker, who has global information about the system and can reason about the choices of others. However, a typical user often lacks the information or ability to make complex predictions regarding future decisions of other users. Rather, users typically employ some simple decision rule, based on their experience.

Accordingly, we proceed to consider a dynamic learning scheme, in which users base their decisions on accumulated statistical data at the server. No prior information is assumed regarding service in or arrivals to the server, nor do users employ game-theoretic considerations to arrive at their decisions. As shall become evident, the Nash equilibrium of the game considered in the previous sections is a stationary point of the learning scheme considered in the present section. Furthermore, we demonstrate that, through the (individual) employment of such a scheme, user choices converge to the Nash equilibrium.

Specifically, we consider a learning scheme which relies on the framework presented in [2]. In the centralized version, considered first, we assume that the server monitors the average service times of users, depending on the queue length at their arrival instants. Newly arriving users have access to this accumulated data, in addition to the queue length at their arrival, and may use it to assess their performance at the server before deciding whether to join.

The system starts at time 0 without any prior data or operating statistics regarding the expected service. For every $t \geq 0$, let $N_t(x)$ denote the number of users who had joined the server at queue length x and already left it by time

t. Let $\hat{V}_t(x)$ denote the (empirical) average service time of these $N_t(x)$ users. A user which arrives at time t may inspect the current vector $\hat{V}_t$.

Consider then a type-i user which arrives at time t, observes X_t users at service, and has to decide whether to join the server. A natural decision rule is:

$$\text{join the system if } \hat{V}_t(X_t) < \theta^i, \text{ and leave otherwise.} \tag{7}$$

The decision rule (7) is completely deterministic – it prescribes a single choice based on the difference between (estimated) service time and the QoS requirement. Such absence of a randomization option complicates the correspondence between the learning algorithm and the Nash equilibrium solution, as the latter may be randomized. Moreover, if the difference between the expected service time and the QoS requirement is small, it stands to reason that some other factors may influence the actual decision. We model these factors by possibly randomized decisions, and consider the following decision rule in place of (7):

$$\text{join the system with probability} \quad S_\epsilon(\theta^i - \hat{V}_t(X_t)) \ , \tag{8}$$

where ϵ is a small positive parameter, and S_ϵ is an increasing function, with $S_\epsilon(x) = 0$ for $x \leq -\epsilon$, and $S_\epsilon(x) = 1$ for $x \geq \epsilon$.

The performance of this learning scheme has been investigated in [6] by way of simulations. The results, some of which are depicted below, demonstrate the convergence of the collective decision rule to the one implied by the Nash equilibrium solution, and consequently the convergence of the system performance measures, such as the average service times or queue length.

The above decision rule relies on *central* statistics, collected and advertised by the system server. Often, experience is accumulated in a more distributed manner, by individuals or groups of users which repeatedly return to the system, and the convergence of such distributed learning schemes is therefore of considerable practical interest. In relation with the present study, we considered a distributed version in which statistical data is *type-specific*, i.e., a user is cognizant only of the experience of its type-peers. Our simulations [6] indicate that the convergence properties of the learning process are retained in this case, although the convergence rate can be slower than in the centralized case due to the dilution in available information.

For illustration, we present here two simulation runs. In the first we apply the basic learning scheme, with centralized statistics, to a system with service rate $\mu = 100$, buffer size $B = 50$, and two types of users with parameters $\theta^1 = 0.1$,

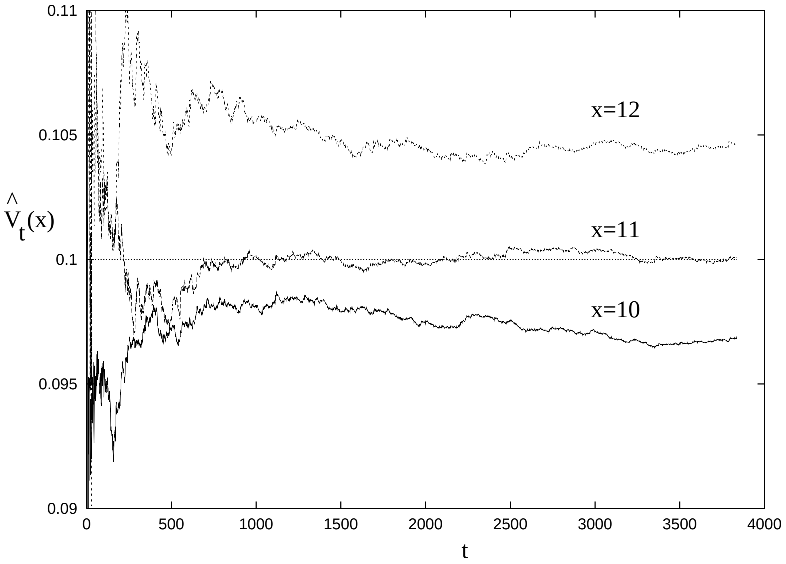

Figure 3. Simulation 1 – empirical service time estimates for queue lengths about the equilibrium threshold of type 1 users

$\lambda^1 = 50$, $\theta^2 = 0.05$ and $\lambda^2 = 100$. The computed equilibrium thresholds are $g^1 = 11.4658$ and $g^2 = 3.64703$. This system was simulated with the (centralized)

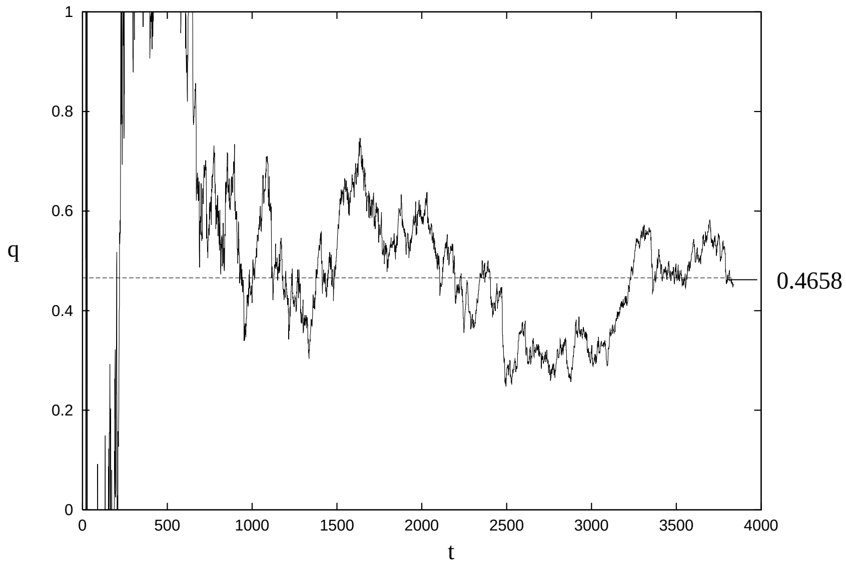

Figure 4. Simulation 1 – the randomized decision probability of type 1 users at their threshold queue length

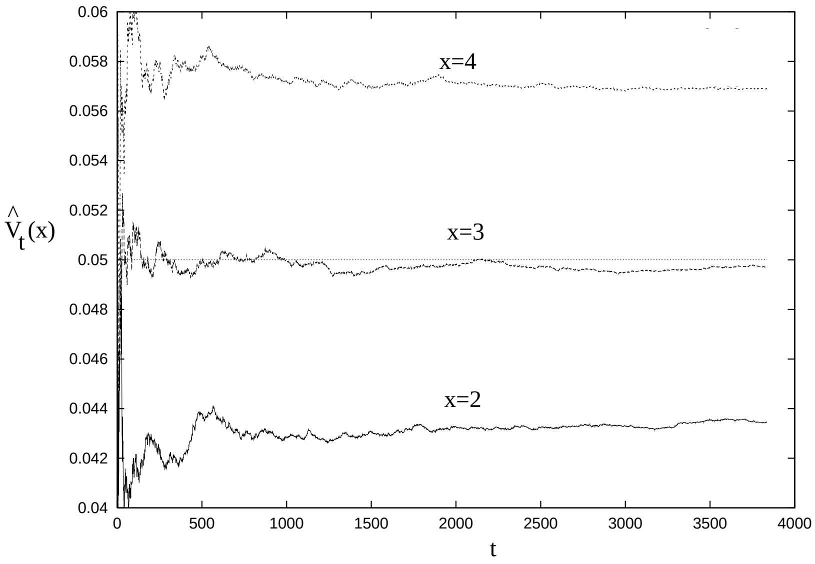

Figure 5. Simulation 1 – empirical service time estimates for queue lengths about the equilibrium threshold of type 2 users

learning scheme (8), taking $\epsilon = 0.001$ and S_ϵ linear. Figure 3 depicts the empirical estimate of the service times $\hat{V}_t(x)$ for the three queue lengths $x = 10, 11, 12$

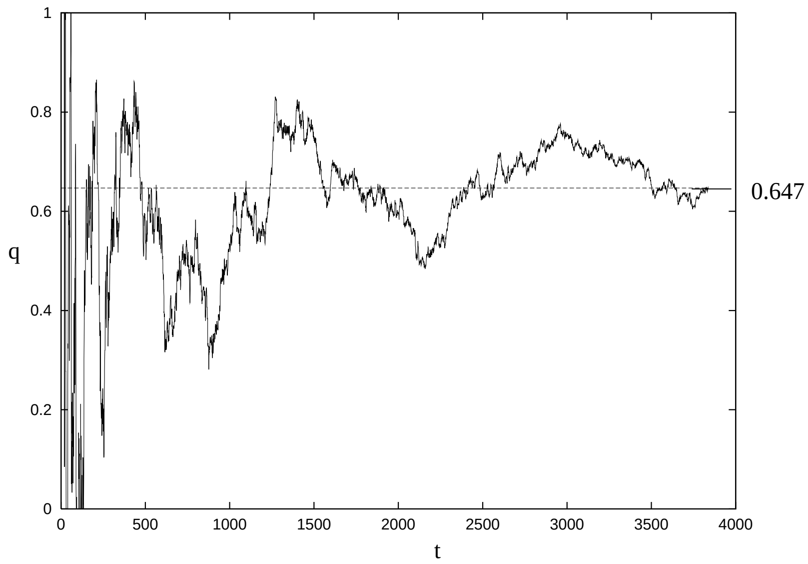

Figure 6. Simulation 1 – the randomized decision probability of type 2 users at their threshold queue length

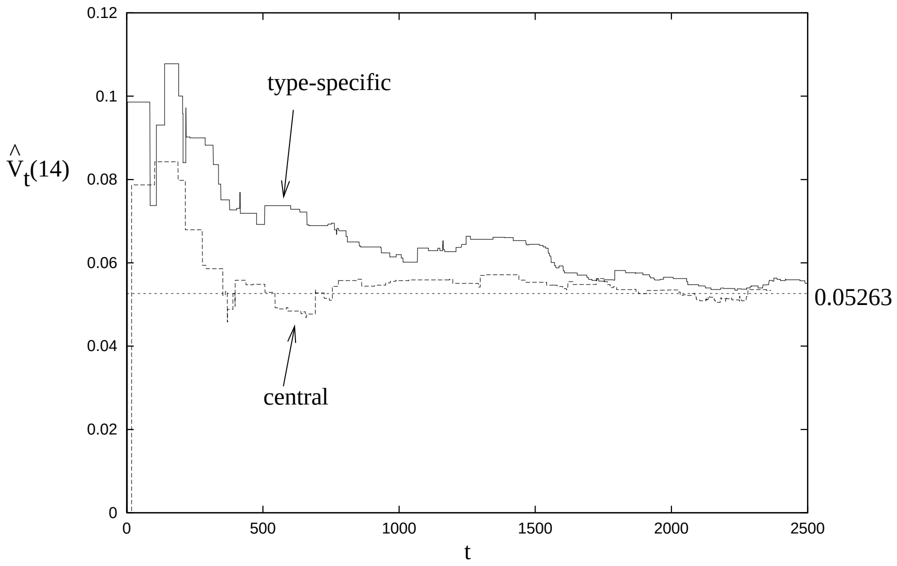

Figure 7. Simulation 2: a comparison of learning rates for centralized and type-specific statistics. Type-specific estimates are based on statistics of the first user type only

about the equilibrium threshold of type 1 users. The QoS parameter $\theta^1 = 0.1$ is also drawn for reference. It may be seen that at queue length $x = 12$, which is above that threshold, the estimate quickly assumes values higher than θ^1, and consequently the decision of type-1 users will be to leave the system when confronted with this queue size. The opposite is true for the queue size $x = 10$, which is below the threshold. For $x = 11$, the randomized equilibrium threshold $g^1 = 11.4658$ implies that $V(11) = \theta^1 = 0.1$; indeed, it may be seen that the estimate $\hat{V}_t(11)$ converges to this value. Figure 4 shows the probability of joining the queue at $x = 11$, which is obtained from $\hat{V}_t(11)$ through (8); it also converges to a close proximity of the equilibrium probability $q^1 = 0.4658$. Figures 5 and 6 show similar results which pertain to type 2 users. Note that the small offset in the convergence of of $\hat{V}_t(3)$ here is just of the required magnitude to induce the equilibrium value of q^2.

The second simulation compares the centralized information learning scheme with the type-specific one. The simulated system here has a service rate of $\mu = 200$, buffer size $B = 50$, and two types of users, with parameters $\theta^1 = \frac{1}{19} \approx 0.05263$, $\lambda^1 = 56$, $\theta^2 = \frac{1}{26} \approx 0.03846$ and $\lambda^2 = 112$. The computed equilibrium thresholds are $g^1 = 14.1425$ and $g^2 = 8.40045$. Figure 7 presents the empirical delays for the first user type at queue size $X = 14$, i.e., about its threshold; as

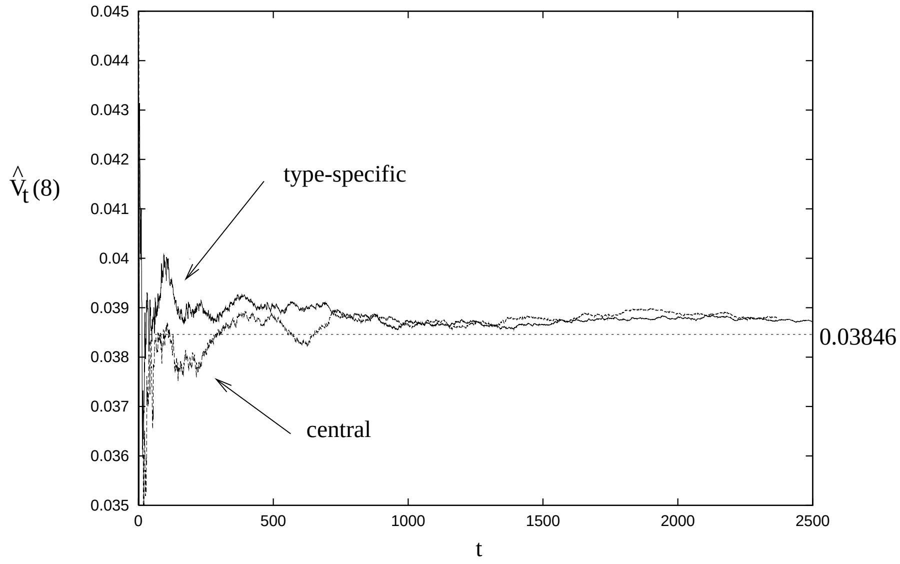

Figure 8. Simulation 2 – type-specific estimates based on second user type statistics

may be seen, with both schemes the empirical values converge to the computed equilibrium value 0.05263, but at a higher rate for the central one. Similarly, Figure 8 presents the results for the second user type at queue size $X = 8$, i.e., about its threshold; here. too, convergence is obtained to the computed value 0.03846 with both schemes and with a somewhat higher rate for the central one.

In [2] convergence has been proved analytically, for the *homogeneous*, i.e. single-type, case, and with central statistics. A slightly modified scheme was analyzed in [7]. These proofs strongly rely on some single-parameter monotonicity properties, which are not available in the heterogeneous case considered here. Convergence analysis for the latter case appears to be considerably more complex.

8. Conclusion

In this paper we considered the behavior of a shared service facility, in which users make individual joining decisions based on their QoS requirements. Users decide whether to accept the available level of service, based on their observation of the current system load. Formulated as an individual decision problem, this setup leads to a dynamic game model, which captures the inherent interaction and conflict among users that share a common resource.

Our analysis established the existence, uniqueness and structure of the Nash equilibrium point. Uniqueness was established under the natural assumption that users with identical QoS requirements will employ the same decision rule. The equilibrium policy profile has a threshold structure, specifying a different threshold for each user type. These findings uniquely characterize the stationary operating point of the system.

The relevance of the Nash equilibrium solution to actual system operation has been demonstrated in a dynamic learning scenario, where users employ simple decision rules based on their accumulated experience. Convergence to the equilibrium point has been observed in both central and distributed learning environments.

We conclude by pointing to some possible directions for further research. In the model considered here, users form a single class in terms of the load they impose on the server and their service priorities. The multi-class case presents an interesting and challenging extension, yet it entails a major difficulty due to the multi-dimensional queue state. Several variants can be formulated, depending on the class structure and the extent of the load information available to an arriving user. Some initial experiments, which involve asymmetric service rates to several user classes, have been conducted and are presented in [6], with encouraging results regarding the convergence of an appropriate learning algorithm.

Convergence analysis of the proposed learning scheme is an open problem in the heterogeneous user case studied here. As has been shown in the homogeneous-user case ([2,7]), Stochastic Approximations provide an appropriate framework for this analysis. Further understanding of structural properties, in particular of the best-response map, will be required in order to tackle the convergence issue in the present case.

References

[1] E. Altman. Flow Control Using the Theory of Zero-Sum Markov Games. *IEEE Transactions on Automatic Control*, 39(4):814–818, April 1994.

[2] E. Altman and N. Shimkin. Individual Equilibrium and Learning in a Processor Sharing System. *Operations Research*, 46(6):776–784, December 1998. Extended version: Technical report CC212, EE Department, Technion, October 1997.

[3] D. Assaf and M. Haviv. Reneging from Processor Sharing Systems and Random Queues. *Mathematics of Operations Research*, 15:129–138, 1990.

[4] T. Basar and G. Oldser. *Dynamic Non Cooperative Game Theory.* Academic Press, London, 1982.

[5] I. Ben-Shahar, A. Orda, and N. Shimkin. Best-Effort Resource Sharing by Users with QoS Requirements. In *Proceedings of the IEEE INFOCOM*, pages 617–624, April 1998.

[6] I. Ben-Shahar. Routing of Individual Users with QoS Requirements. M.Sc. Thesis, Department of Electrical Engineering, Technion, 1998.

[7] R. Buche and H. J. Kushner. Stochastic Approximation and User Adaptation in a Competitive Resource Sharing System. Preprint, Mar. 1998. Brown University.

[8] A. A. Economides and J. A. Silvester. Multi-Objective Routing in Integrated Services Networks: A Game Theory Approach. In *Proceedings of the IEEE INFOCOM'91*, pages 1220–1225, 1991.

[9] M. Haviv. Stable Srategies for Processor Sharing Systems. *European Journal of Operations Research*, 52:1051–1066, 1991.

[10] M.-T. T. Hsiao and A. A. Lazar. Optimal Decentralized Flow Control of Markovian Queueing Networks with Multiple Controllers. *Performance Evaluation*, 13(3):181–204, 1991.

[11] Y. A. Korilis and A. A. Lazar. On the Existence of Equilibria in Noncooperative Optimal Flow Control. *Journal of the ACM*, 42(3):584–613, May 1995.

[12] A. A. Lazar, A. Orda, and D. E. Pendarakis. Virtual Path Bandwidth Allocation in Multi-User Networks. *IEEE/ACM Transactions on Networking*, 5(6):861–871, December 1997.

[13] A. Mandelbaum and N. Shimkin. A Model for Rational Abandonments from Invisible Queues. Submitted to *Queueing Systems*, 1999.

[14] R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, Massachusetts, 1991.

[15] P. Naor. On the Regulation of Queue Size by Levying Tolls. *Econometrica*, 37:15–24, 1969.

[16] A. Orda, R. Rom, and N. Shimkin. Competitive Routing in Multiuser Communication Networks. *IEEE/ACM Transactions on Networking*, 1(5):510–521, Oct. 1993.

[17] S. Stidham. Optimal Control of Admission to a Queuing System. *IEEE Transactions on Automatic Control*, 30:705–713, 1985.

[18] U. Yechiali. On Optimal Balking Rules and Tool Charges in the GI/M/1 Queuing Process. *Operations Research*, 19(2), 1971.

[19] U. Yechiali. Customers' Optimal Joining Rules for the GI/M/s Queue. *Management Science*, 18:434–443, 1972.

[20] Z. Zhang and C. Douligeris. Convergence of Synchronous and Asynchronous Greedy Algorithms in a Multiclass Telecommunications Environment. *IEEE Transactions on Communications*, 40(8):1277–1281, Aug. 1992.