FLOW CONTROL USING THE THEORY OF ZERO SUM MARKOV GAMES

Eitan ALTMAN*
INRIA
Centre Sophia Antipolis
06565 Valbonne Cedex, France
altman@martingale.inria.fr, tel. 93 65 76 73

September 1992

Abstract

We consider the problem of dynamic flow control of arriving packets into an infinite buffer. The service rate may depend on the state of the system, may change in time and is unknown to the controller. The goal of the controller is to design an efficient policy which guarantees the best performance under the worst service conditions. The cost is composed of a holding cost, a cost for rejecting customers (packets) and a cost that depends on the quality of the service. The problem is studied in the framework of zero-sum Markov games, and a value iteration algorithm is used to solve it. We show that there exists an optimal stationary policy (such that the decisions depend only on the actual number of customers in the queue); it is of a threshold type, and it uses randomization in at most one state.

1 Introduction

Game theoretical methods seem to be quite promising in the control of queueing systems in general and Telecommunication systems in particular. In the latter, one often encounters situations where different users may have different objectives. Typically, each user would like to optimize some performance measures related to his traffic, such as minimizing delays, maximizing his throughput and minimizing blocking probabilities. The situation of game arises when the different users cannot coordinate their actions, and the problem cannot be reduced to a single controller case. This may

*The work of this author was supported by the Chateaubriand fellowship from the french embassy in Israel

happen either because the control is decentralized due to the nature of the network, or simply because any user may be tempted to benefit on the expense of the others by choosing an individual good policy.

There has been some work in applying game theory to control of queueing networks and control in telecommunications. Among these are several papers that consider static control problems and thus use static games: Marchand [6] considers optimal input control to a queueing system; Hsiao and Lazar [2], Douligeris and Mazumdar [1], consider the problem of optimal flow control in a multiclass telecommunications environment; Mason and Girard [7, 8] consider routing models; Shenker [9] studies the problem of optimal service allocation to several users. Kalai and Zemel [3] consider a problem in networks that yield a cooperative game.

The main aim of this paper is to show how methods based on Markov games (also known as stochastic games) may be applied to **dynamic control in queueing networks** under a min-max criterion. We consider a control problem with a single controller, who controls dynamically the flow of arriving packets into an infinite buffer. We assume that the service rate may depend on the state of the system, and may change in time in a way that is unknown to the controller. The goal of the controller is to design an efficient policy which guarantees the best performance under the worst service conditions. The cost is composed of a holding cost, a cost for rejecting customers and a cost that depends on the quality of the service. Since nothing is known on the behavior of the server, a pesimistic assumption is used that regards the server as a player in a zero-sum game who is playing against the flow controller. A value iteration algorithm is used to solve this problem. We show that there exists an optimal stationary policy (such that the decisions depend only on the actual number of customers in the queue); it is of a threshold type, and it uses randomization in at most one state.

The structure of the paper is as follows: in Section 2 we describe the model. Then basic tools for solving the problem are described in Section 3, and the optimal policy is derived in Section 4.

2 The Model

Considered is a discrete-time single-server queue with an infinite buffer. We assume that at most one customer may join the system in a time slot. This arrival (if any) is assumed to occur at the beginning of the time slot (synchronized arrivals). At the beginning of each time slot, the flow control mechanism, called player 2, chooses in the set $\mathbf{A} := \{a_1, a_2\}$, $0 < a_2 \leq a_1 \leq 1$, the probability of having one arrival in this time slot. Therefore, if action a_i is chosen at time t then a customer will enter the system in $[t, t+1)$ with the probability a_i, $i = 1, 2$.

At the end of each slot, a successful service of a customer occurs, if the queue is not empty, with probability $b \in \mathbf{B} = \{b_1, b_2\}$, where $1 > b_1 \geq b_2 \geq 0$. (If the service fails the customer remains in the queue). The value of b, which may represent the quality of service, may change in each time slot, and is not known to player 2. The objective of player 2, to be described below, is to find a best strategy under the worst case service conditions. We model the system as a zero-sum Markov game, where player 1 controls the service quality.

Actions a and b are assumed to be taken independently, based on the information on the current state as well as the information of all past states and actions of both players.

We assume that a customer that enters an empty system may leave the system (with the probability b) at the end of this same time slot.

Let X_t denote the number of customers in the system at time t, $t \in \mathbb{N}$, and B_t and A_t denote the actions of players 1 and 2 respectively. The state space is denoted by $\mathbf{X} = \{0, 1, ...\}$. Let $M(Y)$ be the set of probability measures on a set Y.

The transition law q is:

$$q(y \mid x; a; b) := \begin{cases} \bar{a}b, & \text{if } x \geq 1,\ y = x - 1; \\ ab + \bar{a}\bar{b}, & \text{if } x \geq 1,\ y = x; \\ a\bar{b}, & \text{if } x \geq 0,\ y = x + 1; \\ 1 - a\bar{b}, & \text{if } y = x = 0; \end{cases}$$

(for any number $\xi \in [0, 1]$, $\bar{\xi} := 1 - \xi$).

We define an immediate payoff

$$C(x, b, a) := c(x) + \gamma\, a + \theta\, b, \tag{2.1}$$

for all $x \in \mathbf{X}$, $a \in \mathbf{A}$ and $b \in \mathbf{B}$. $C(x, b, a)$ is the cost that player 2 pays to player 1 when the state is x, and the actions of the players are b and a. C generalizes a cost frequently encountered in the literature on flow control models. In (2.1) $c(x)$ is any real-valued *nondecreasing convex* function on $\mathbb{N}$ and γ and θ are arbitrary real constants. We shall assume that c is nonnegative and it is upper bounded by some polynome; $\gamma \leq 0$ and $\theta \geq 0$. $c(x)$ can be interpreted as a holding cost per unit time, γ as a reward related to the acceptance of an incoming customer, and θ as a cost per quality of service.

Let U (V) be the class of policies (strategies) of player 1 (player 2 resp.). A policy $u \in U$ $(v \in V)$ is a sequence $u = (u_0, u_1, ...)$ $(v = (v_0, v_1, ...)$ resp.) where u_n (resp. v_n) is a conditional probability on $\mathbf{B}$ (resp. $\mathbf{A}$) given the history of all states and actions of both players as well as the current state (see definition in [10]).

Let ξ be fixed number in $[0,1)$. Define

$$V_\xi(x,u,v) := E^{u,v}\left[\sum_{t=0}^{\infty} \xi^t\, C(X_t,B_t,A_t) \,|\, X_0 = x\right].$$

Define the following problem ($\mathbf{Q}_\xi$): find v that achieves

$$V_\xi(x) := \inf_{v\in V} \sup_{u\in U} V_\xi(x,u,v), \qquad \forall x \in \mathbf{X}.$$

$V_\xi(x)$ is called the ξ-discounted *value* of the game, due to the saddle point property described in Proposition 3.1 (ii). A policy of a player is said to be stationary if it depends only on the current state (and not on the previous history nor on the time). For stationary policies u and v of the two players we shall denote (with some abuse of notation) $u_x = \{u_x(1), u_x(2)\} \in M(\mathbf{B})$, where $u_x(i)$ is the probability that the first player chooses b_i at state x; and $v_x = \{v_x(1), v_x(2)\} \in M(\mathbf{A})$, where $v_x(i)$ is the probability that the second player chooses a_i at state x.

It can be shown that the problem $\mathbf{Q}_\xi$ becomes trivial when $\gamma \geq 0$. In that case, the optimal action of player 2 is always to choose a_2. Similarly, if $\theta \leq 0$ then it is optimal for player 1 always to choose b_2.

Remark: One could also consider the problem where $\mathbf{A}$ is any subset of the interval $\mathbf{A} = [a_2, a_1]$. However, it can be shown that any choice of an action a that satisfies $a_2 < a < a_1$ is equivalent to randomly choosing between a_2 and a_1 with some probabilities. Therefore, the optimal policy for the problem that we study is also optimal for the problem with $\mathbf{A} \subset [a_2, a_1]$. A similar argument holds for the set $\mathbf{B}$.

3 Preliminary Results

Let $\mathcal{K}$ be the set of all real-valued functions $f : \{-1\} \cup \mathbf{X} \to \mathbb{R}$ that are polynomially bounded and satisfy $f(-1) = f(0)$. Let the operator $R : \mathbf{X} \times \{1,2\} \times \{1,2\} \times \mathcal{K} \to \mathbb{R}$ be defined as

$$R(x,i,j,f) := E\left[f(X_{t+1})|X_t = x, B_t = b_i, A_t = a_j\right]$$

we get for $x \in \mathbf{X}$:

$$R(x,i,j,f) = \bar{a}_j b_i f(x-1) + (a_j b_i + \bar{a}_j \bar{b}_i) f(x) + a_j \bar{b}_i f(x+1).$$

Define the operator $S : (\{-1\} \cup \mathbf{X}) \times \{1,2\} \times \{1,2\} \times \mathcal{K} \to \mathbb{R}$ as

$$S(x,i,j,f) := C(x,b_i,a_j) + \xi R(x,i,j,f), \qquad x \geq 0,$$

and let $S(-1,i,j,f) = S(0,i,j,f)$. Let $S(x,f)$ denote the matrix whose entries are $S(x,i,j,f)$.

For any function $D : \mathbf{A} \times \mathbf{B} \to \mathbb{R}$, $\alpha \in M(\mathbf{A})$ and $\beta \in M(\mathbf{B})$ define

$$D(\beta,\alpha) = \beta D \alpha = \sum_{i=1}^{2}\sum_{j=1}^{2} \beta_i \alpha_j D(b_i, a_j)$$

The value of the matrix game D is defined as $val(D) := \sup_{\beta \in M(\mathbf{B})} \inf_{\alpha \in M(\mathbf{A}_x)} \beta D \alpha$. It is known to satisfy $val(D) = \inf_{\alpha \in M(\mathbf{A}_x)} \sup_{\beta \in M(\mathbf{B})} \beta D \alpha$, and there are probability measures $\beta^* \in M(\mathbf{B})$, $\alpha^* \in M(\mathbf{A}_x)$ such that

$$val(D) = \inf_{\alpha \in M(\mathbf{A}_x)} \beta^* D \alpha = \sup_{\beta \in M(\mathbf{B})} \beta D \alpha^* = \beta^* D \alpha^*$$

β^* and α^* are said to be optimal for the matrix game D.

Let $T_\xi : \mathcal{K} \to \mathcal{K}$ be the DP operator associated with $\mathbf{Q}_\xi$: $T_\xi f(x) := val\, S(x,f)$, $x \in \mathbf{X}$. We shall use the following tools for solving $\mathbf{Q}_\xi$:

Proposition 3.1 *(i) V_ξ satisfies*

$$V_\xi(x) = T_\xi V_\xi(x) \tag{3.1}$$

(ii) Let (u^*, v^*) be the stationary policies of player 1 and 2 respectively that achieve the sup and the inf in the matrix game $S(x, V_\xi)$ (i.e. the right side of 3.1). Then (u^*, v^*) are optimal for $\mathbf{Q}_\xi$ and

$$\begin{aligned} \sup_{u \in U} \inf_{v \in V} V_\xi(x,u,v) &= \inf_{v \in V} \sup_{u \in U} V_\xi(x,u,v) = \inf_{v \in V} V_\xi(x,u^*,v) \\ &= \sup_{u \in U} V_\xi(x,u,v^*) = V_\xi(x,u^*,v^*) = V_\xi(x). \end{aligned}$$

(iii) For every $f \in \mathcal{K}$, $\lim_{n\to\infty} T_\xi^n f = V_\xi$.

Proof. The theorem follows from Wessels [10]. We show below that the contraction assumptions there are indeed satisfied. We begin by establishing Assumption 2.1, which says that there exists some bounding function $\mu^{-1}(x) > 0$ under which $\xi R(x,i,j,\mu^{-1})$ is a contraction for all x, i, j, i.e. there exists some α, $0 < \alpha < 1$ such that

$$\xi |R(x,i,j,\mu^{-1})| \le \alpha \mu^{-1}(x), \quad \forall x, i, j. \tag{3.2}$$

Choose some $\hat{\xi}$, $\xi < \hat{\xi} < 1$. In order to establish (3.2) it suffices to show that

$$\hat{\xi} |R(x,i,j,\mu^{-1})| \le \mu^{-1}(x), \quad \forall x, i, j$$

Fix some $i, j \in \{1,2\}$ and let $a = a_j$ and $b = b_i$. Define $r := 1 + (\hat{\xi}^{-1} - 1)/a\bar{b}$ and let $\mu(x) := r^{-x}$. It follows that

$$\hat{\xi}R(x,i,j,\mu^{-1})(x) - \mu^{-1}(x) =$$

$$\begin{cases} \left[a\bar{b}r^2 + (ab + \bar{a}\bar{b} - \hat{\xi}^{-1})r + \bar{a}b\right]\hat{\xi}r^{x-1} = -\hat{\xi}r^{x-1}\bar{a}b(r-1) < 0, & x > 0 \\ \hat{\xi}[a\bar{b}r + (1 - a\bar{b} - \hat{\xi}^{-1})] = 0, & x = 0 \end{cases}$$

which establishes Assumption 2.1 in Wessels [10]. Next we show that Assumption 2.2 holds too. The assumption states that the cost function is finite under the norm induced by the bounding function μ. More precisely, we need to show that $\sup_{x,b,a}\{|C(x,b,a)|\mu(x)\} < \infty$. The latter indeed follows from the fact that $c(x)$ is polynomially bounded whereas $\mu^{-1}(x)$ grows exponentially fast (since $r > 1$). This establishes Assumption 2.2 in [10]. The Theorem finally follows from Wessels [10] Theorem 3.1 and Section 4 there. ∎

4 The Optimal Policy

Proposition 3.1 is the main tool to establish the structure of the optimal policy, given in Theorem 4.2 below. It is used in the following way. We first guess the structure of an optimal stationary policy. We then identify sufficient conditions that V_ξ should possess in order to ensure this structure of the optimal policy; the connection between the optimal policy and V_ξ is given in Proposition 3.1 (ii). We then use the "value iteration" technique, based on Proposition 3.1 (iii) to show that V_ξ indeed possess the required properties. This is done iteratively: we show that there exists some function f that possesses them. Then we show that if a function f has these properties then so does $T_\xi f$, and hence also $T_\xi^n f$ for all $n \in \mathbb{N}$. Finally we show that the latter implies also that the limit as n goes to infinity also possesses these properties, and hence by Proposition 3.1 (iii) V_ξ possesses the properties as well. We thus begin the analysis by listing some properties that a function may possess.

We shall say that $g : (\{-1\} \cup \mathbf{X}) \times \{1,2\} \times \{1,2\} \to \mathbb{R}$ satisfies:

J1 if $g(x+1,i,2) - g(x+1,i,1) \le g(x,j,2) - g(x,j,1)$ for $x \in \mathbf{X}$, $i,j = 1,2$.

J2 if $g(x+1,1,i) - g(x+1,2,i) \le g(x,1,j) - g(x,2,j)$ for $x \in \mathbf{X}$, $i,j = 1,2$.

J3 if $g(x,1,1) - g(x,2,1) \le g(x,1,2) - g(x,2,2)$ for $x \in \mathbf{X}$.

J4 if $g(x+2,i,j) - g(x+1,i,k) \ge g(x+1,i,j) - g(x,i,k)$ for $x \in \mathbf{X}$, $i,j,k = 1,2$.

We shall say that $f \in \mathcal{K}$ satisfies:

IC if $f(x)$ is convex increasing in x; equivalently, for any $x \in \mathbf{X} \cup \{-1\}$,

$$f(x+2) - f(x+1) \geq f(x+1) - f(x). \tag{4.1}$$

(Indeed, since $f(-1) = f(0)$ **IC** implies that f is increasing).

Let $(u(f), v(f))$ be the stationary policies that achieve the sup and the inf in the matrix game $S(x, f)$ for all $x \in \mathbf{X}$. For $f = V_\xi$ we have in particular $(u(f), v(f)) = (u^*, v^*)$. Any $(u(f), v(f))$ for which f satisfies **IC** have the following threshold structure, where randomization occurs in at most one state:

Theorem 4.1 *Assume that f satisfies* **IC**. *Then there exist $m_u, m_v \in \mathbf{X}$ (that depend on f) and stationary policies u, v, such that*

$$u_x(f) = \begin{cases} (1,0) & \text{if } x < m_u \\ (q_u, \bar{q}_u) & \text{if } x = m_u \text{ and } m_u = m_v \\ (0,1) & \text{if } x = m_u \text{ and } m_u \neq m_v \\ (0,1) & \text{if } x > m_u \end{cases} \qquad v_x(f) = \begin{cases} (1,0) & \text{if } x < m_v \\ (q_v, \bar{q}_v) & \text{if } x = m_v \text{ and } m_u = m_v \\ (0,1) & \text{if } x = m_v \text{ and } m_u \neq m_v \\ (0,1) & \text{if } x > m_v \end{cases}$$

where $q_u, q_v \in [0,1]$ are some constants.

The proof of Theorem 4.1 relies upon the following technical lemmas:

Lemma 4.1 *Let $h : \mathbf{X} \cup \{-1\} \to \mathbb{R}$ be a nondecreasing function. Let $\zeta_1, \zeta_2 \in [0,1]$. Then, for all $0 \leq x$,*

$$F(x) := \zeta_2 h(x+1) + \bar{\zeta}_2 h(x) - \zeta_1 h(x) - \bar{\zeta}_1 h(x-1) \geq 0 \tag{4.2}$$

Proof.

$$F(x) \geq h(x) - \zeta_1 h(x) - \bar{\zeta}_1 h(x-1) = \bar{\zeta}_1 [h(x) - h(x-1)] \geq 0$$

■

Lemma 4.2 *Assume that f satisfies* **IC**. *Then, $S(\bullet, f)$ satisfies* **J1**- **J4**, *and for any $i, j \in \{1,2\}$, $S(\bullet, i, j, f)$ satisfies* **IC**.

Proof. We begin by **J1**. For $0 \leq x$ and any $i, j = 1, 2$,

$$\begin{aligned} & S(x+1,i,1,f) - S(x+1,i,2,f) - [S(x,j,1,f) - S(x,j,2,f)] \\ = \;& \xi(a_1 - a_2)\Big(b_i[f(x+1) - f(x)] + \bar{b}_i[f(x+2) - f(x+1)] \\ & - [b_j[f(x) - f(x-1)] + \bar{b}_j[f(x+1) - f(x)]]\Big) \geq 0 \end{aligned}$$

where the last inequality follows from Lemma 4.1 with $\zeta_2 = \bar{b}_i$, $\zeta_1 = \bar{b}_j$, $h(x) = f(x+1) - f(x)$, and since f satisfies **IC**.

Next we check **J2**. For $0 \leq x$ and any $i, j = 1, 2$,

$$\begin{aligned} & S(x+1,1,i,f) - S(x+1,2,i,f) - [S(x,1,j,f) - S(x,2,j,f)] \\ = \;& -\xi(b_1 - b_2)\Big(\bar{a}_i[f(x+1) - f(x)] + a_i[f(x+2) - f(x+1)] \\ & - [\bar{a}_j[f(x) - f(x-1)] + a_j[f(x+1) - f(x)]]\Big) \leq 0 \end{aligned}$$

where the last inequality follows from Lemma 4.1 with $\zeta_2 = a_i$, $\zeta_1 = a_j$ and $h(x) = f(x+1) - f(x)$, and since f satisfies **IC**.

Next we check **J3**. For $0 \leq x$,

$$\begin{aligned} & S(x,1,1,f) - S(x,2,1,f) - [S(x,1,2,f) - S(x,2,2,f)] \\ = \;& -\xi(b_1 - b_2)(a_1 - a_2)(f(x+1) - f(x) - f(x) + f(x-1)) \leq 0 \end{aligned}$$

since f satisfies **IC**.

Next we check **J4**. For $0 \leq x$, $i, j, k = 1, 2$,

$$\begin{aligned} & S(x+2,i,j,f) - S(x+1,i,k,f) - [S(x+1,i,j,f) - S(x,i,k,f)] \\ = \;& c(x+2) - c(x+1) - c(x+1) + c(x) + \\ & \xi\Big\{ b_i\left[a_j f(x+2) + \bar{a}_j f(x+1) - a_k f(x+1) - \bar{a}_k f(x)\right] \\ & +\bar{b}_i\left[a_j f(x+3) + \bar{a}_j f(x+2) - a_k f(x+2) - \bar{a}_k f(x+1)\right] \\ & -b_i\left[a_j f(x+1) + \bar{a}_j f(x) - a_k f(x) - \bar{a}_k f(x-1)\right] \\ & -\bar{b}_i\left[a_j f(x+2) + \bar{a}_j f(x+1) - a_k f(x+1) - \bar{a}_k f(x)\right]\Big\} \\ \geq \;& \xi\Big\{ b_i\left[a_j(f(x+2) - f(x+1)) + \bar{a}_j(f(x+1) - f(x))\right. \end{aligned}$$

$$- \ a_k(f(x+1)-f(x)) - \bar{a}_k(f(x)-f(x-1))] \Big\}$$

$$+\xi \Big\{ \bar{b}_i \left[a_j(f(x+3)-f(x+2)) \ + \bar{a}_j(f(x+2)-f(x+1))\right.$$

$$\left. -a_k(f(x+2)-f(x+1)) - \bar{a}_k(f(x+1)-f(x))\right] \Big\} \geq 0$$

which follows by applying twice Lemma 4.1 with $\zeta_2 = a_j$ and $\zeta_1 = a_k$; once with $h(x) = f(x+1) - f(x)$ for the term in the first curly brackets, and once with $h(x) = f(x+2) - f(x-1)$ for the term in the second curly brackets.

That $S(\bullet, i, j, f)$ satisfies **IC** for $x \in \mathbf{X}$ follows from the fact that $S(\bullet, f)$ satisfies **J4**. For $x = -1$ we note that $S(0, i, j, f) = S(-1, i, j, f)$ by definition, and hence it suffices to show that $S(1, i, j, f) - S(0, i, j, f) \geq 0$. This follows easily from the fact that f is increasing. ∎

Lemma 4.3 *Assume that f satisfies* **IC** *and define*

$$l_i^J := \min_{y \in \mathbf{X}} \{S(y, i, 2, f) \leq S(y, i, 1, f)\}, \ i = 1, 2$$

$$l_j^I := \min_{y \in \mathbf{X}} \{S(y, 1, j, f) \leq S(y, 2, j, f)\}, \ j = 1, 2$$

Then (i) $S(x, i, 1, f) \geq S(x, i, 2, f)$ iff $x \geq l_i^J$, $i = 1, 2$; $S(x, 1, j, f) \leq S(x, 2, j, f)$ iff $x \geq l_j^I$, $j = 1, 2$. (ii) l_i^j satisfy:

$$l_2^J \leq l_1^J \leq l_2^J + 1 \tag{4.3}$$

$$l_1^I \leq l_2^I \leq l_1^I + 1 \tag{4.4}$$

Proof. (i) follows since $S(\bullet, f)$ satisfies **J1** and **J2**. Since $S(\bullet, f)$ satisfies **J3**, by Lemma 4.2, it follows that

$$\{S(x, 1, 1, f) \geq S(x, 1, 2, f)\} \Rightarrow \{S(x, 2, 1, f) \geq S(x, 2, 2, f)\}$$

and hence the left inequality in (4.3). Moreover,

$$\{S(x, 1, 2, f) \leq S(x, 2, 2, f)\} \Rightarrow \{S(x, 1, 1, f) \leq S(x, 2, 1, f)\}$$

and hence the left inequality in (4.4). Since $S(\bullet, f)$ satisfies **J1**, by Lemma 4.2, it follows that

$$\begin{aligned} \{x \geq l_2^J\} \ &\Rightarrow \ \{S(x, 2, 1, f) \geq S(x, 2, 2, f)\} \\ &\Rightarrow \ \{S(x+1, 1, 1, f) \geq S(x+1, 1, 2, f)\} \Rightarrow \{x+1 \geq l_1^J\} \end{aligned}$$

and hence the right inequality in (4.3). Finally, since $S(\bullet, f)$ satisfies **J2**, by Lemma 4.2, it follows that

$$\begin{aligned}\{x \geq l_1^I\} \;&\Rightarrow\; \{S(x,1,1,f) \leq S(x,2,1,f)\}\\ &\Rightarrow\; \{S(x+1,1,2,f) \leq S(x+1,2,2,f)\} \Rightarrow \{x+1 \geq l_2^I\}\end{aligned}$$

and hence the right inequality in (4.4). This establishes (ii). ∎

Proof of Theorem 4.1: It follows from Lemma 4.3 that
(i) $S(x,i,1,f) \geq S(x,i,2,f)$ for $x > l_2^J$ and any $i \in \{1,2\}$;
(ii) $S(x,i,1,f) < S(x,i,2,f)$ for $x < l_2^J$ and any $i \in \{1,2\}$;
(iii) $S(x,2,j,f) \geq S(x,1,j,f)$ for $x > l_1^I$ and any $j \in \{1,2\}$;
(iv) $S(x,2,j,f) < S(x,1,j,f)$ for $x < l_1^I$ and any $j \in \{1,2\}$;

To obtain $valS(x,f)$, we note the following: if $S(x,i,1,f) \geq S(x,i,2,f)$ for both $i = 1,2$, then player 2 has a dominant policy, i.e. in state x it is optimal for him to choose action a_2 (w.p.1). If $S(x,i,1,f) \leq S(x,i,2,f)$ for both $i = 1,2$, then the dominant policy of player 2 when in state x is to choose action a_1. If $S(x,1,j,f) \leq S(x,2,j,f)$ for both $j = 1,2$, then the dominant policy of player 1 when in state x is to choose action b_2, and if $S(x,1,j,f) \geq S(x,2,j,f)$ for both $j = 1,2$, then the dominant policy of player 1 when in state x is to choose action b_1. If player 1 has a dominant policy of choosing action b_i in state x, then (i) if $S(x,i,2,f) \leq S(x,i,1,f)$ then the optimal policy of player 2 in state x is a_2, and $valS(x,f) = S(x,i,2,f)$; (ii) if $S(x,i,1,f) \leq S(x,i,2,f)$ then the optimal policy of player 2 in state x is a_1, and $valS(x,f) = S(x,i,1,f)$. If player 2 has a dominant policy of choosing action a_j in state x, then (i) if $S(x,1,j,f) \leq S(x,2,j,f)$ then the optimal policy of player 1 in state x is b_2, and $valS(x,f) = S(x,2,j,f)$; (ii) if $S(x,2,j,f) \leq S(x,1,j,f)$ then the optimal policy of player 1 in state x is b_1, and $valS(x,f) = S(x,1,1,f)$.

we consider the following three cases.
Case 1: $l_1^I < l_2^J$ (which implies that $l_2^I \leq l_1^J$). Then

$$valS(x,f) = \begin{cases} S(x,1,1,f) & x < l_1^I \\ S(x,2,1,f) & l_1^I \leq x < l_2^J \\ S(x,2,2,f) & l_2^J \leq x \end{cases}$$

We set , $m_v = l_2^J$, $m_u = l_1^I$.
Case 2: $l_2^I > l_1^J$ (which implies that $l_1^I \geq l_1^J$). Then

$$valS(x,f) = \begin{cases} S(x,1,1,f) & x < l_1^J \\ S(x,1,2,f) & l_1^J \leq x < l_2^I \\ S(x,2,2,f) & l_2^I \leq x \end{cases}$$

We set $m_v = l_1^J$, $m_u = l_2^I$.

Case 3: $l_2^I \leq l_1^J$, and $l_1^I \geq l_2^J$. Then we set $m_u = m_v = l_1^I$. There are constants $q_u, q_v \in [0,1]$, such that

$$valS(x,f) = \begin{cases} S(x,1,1,f) & x < l_1^I \\ S(x,u_x,v_x,f) & x = l_1^I \\ S(x,2,2,f) & l_1^I < x \end{cases}$$

where $u_{l_1^I} = \{q_u, \bar{q}_u\}$ and $v_{l_1^I} = \{q_v, \bar{q}_v\}$. q_u and q_v are the optimal solution of the matrix game $S(l_1^I, f)$. Note that in fact randomization occurs only in Case 3 and only if $l_1^I = l_2^J < l_2^I = l_1^J$. The other possibilities in case 3 (for which $q_u = q_v = 0$) are (i) $l_1^I = l_2^I = l_1^J = l_2^J + 1$; (ii) $l_1^I = l_2^I = l_2^J = l_1^J - 1$ and (iii) $l_1^I = l_2^I = l_2^J = l_1^J$. ∎

Lemma 4.4 *Assume that f satisfies* **IC**. *Then $T_\xi f$ satisfies* **IC**.

Proof. It follows from Wessels [10] Section 4 that $T_\xi f \in \mathcal{K}$. $valS(\bullet, f)$ satisfies **IC** since

$$\begin{aligned} & valS(x+2,f) - valS(x+1,f) - [valS(x+1,f) - valS(x,f)] \\ \geq\ & S(x+2, u_{x+1}(f), v_{x+2}(f), f) - S(x+1, u_{x+1}(f), v_x(f), f) \\ & -[S(x+1, u_{x+1}(f), v_{x+2}(f), f) - S(x, u_{x+1}(f), v_x(f), f)] \geq 0 \end{aligned}$$

where the last inequality follows since $S(x,f)$ satisfies **J4**. ∎

We are now ready to prove the main result.

Theorem 4.2 *The optimal value V_ξ of problem ($\mathbf{Q}_\xi$) satisfies* **IC**, *and the optimal policies (u^*, v^*) have a threshold structure described by Theorem 4.1.*

Proof. Choose $f(x) = 0, \forall x \in \mathbf{X} \cup \{-1\}$. It clearly satisfies **IC**. Hence by repeated application of Lemma 4.4, it follows that $T_\xi^n f$ satisfies **IC** $n = 1, 2, ...$; moreover, $\lim_{n\to\infty} T_\xi^n f$ satisfies **IC**. Hence by Proposition 3.1 (iii), V_ξ satisfies **IC**, and the proof is established by applying Theorem 4.1 and Proposition 3.1 (ii). ∎

Acknowledgement: I wish to thank Prof. Jerzy Filar for many fruitful discussions.

References

[1] C. Douligeris and R. Mazumdar, "On Pareto Optimal Flow Control in an Integrated Environment", *Proc. of the 25th Allerton Conference* Univ. of Illinois, Urbana, Oct. 1987.

[2] M. T. Hsiao and A. A. Lazar, "Optimal Decentralized Flow Control of Markovian queueing Networks with Multiple Controllers", *CTR Technical Report*, CUCTR-TR-19, Columbia University, 1986.

[3] E. Kalai and E. Zemel, "Generalized Network Problems Yielding Totally Balanced Games", *Operations Research* **30** No. 5, pp. 998-1008, 1982.

[4] Lazar A. "The Throughput Time Delay Function of an M/M/1 Queue", *IEEE Trans. on Info. Theory*, **IT-29** No. 6, pp. 914-918, 1981.

[5] Lazar A. "Optimal flow control of a class of queuing networks in equilibrium", *IEEE Trans. Auto. Cont.*, Vol 28 no. 11, pp. 1001-1007, 1983.

[6] M. Marchand, "Priority Pricing", *Management Science*, **20**, No. 7, pp. 1131-11400, 1974.

[7] L. Mason, "Flow, Routing Patterns and Algorithms for Store-and-Forward Networks", *J. of Large Scale Systems*, **8**, pp. 187-209, 1985.

[8] L. Mason and A. Girard, "Control Techniques and Performance Models in Circuit Switched Networks", *Proc IEEE CDC*, Orlando, Fl., 1982.

[9] S. Shenker, "Efficient Network Allocation with Selfish Users", *Proceeding of the 14th IFIP WG 7.3 International Symposium on Computer Performance Modeling*, Edinburgh, pp. 279-285, 1990.

[10] J. Wessels, "Markov games with unbounded rewards", *Memorandum COSOR* 76-05, Department of Mathematics, Einthoven University of Technology, 1976, appeared in *Dynamische Optimierung*, M. Schäl editor, Bonner Mathematische Schriftenn, Nr. 98, Bonn, 1977.