Flow control using the theory of zero sum Markov games

Eitan Altman

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

January 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 case 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.

Keywords: Dynamic flow control, Markov zero-sum game, control of queueing networks.

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

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

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 arrizes when the different users cannot coordinate their actions, and the problem cannot be reduced to a single controller case. This may 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 [8] considers optimal input control to a queueing system; Hsiao and Lazar [4], Douligeris and Mazumdar [1, 2, 3], Mazumadar, Mason and Douligeris [11] consider the problem of optimal flow control in a multiclass telecommunications environment; Mason and Girard [9, 10] consider routing models; Shenker [12] studies the problem of optimal service allocation to several users. Kalai and Zemel [5] 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**. Two types of problems may require game theoretical solution methods. The first is control under worst case criterion. Any user, say user A, regards all the other users as one super-user. The aim of A is to guarantee the best performance under the worst case behavior of the super-user. This kind of problems may be solved using zero-sum Markov games. The second type of problems is when the users try to find strategies that leads to Nash equilibrium. In that case every user has some strategy (which may be more profitable than the one obtained from the worst-case criterion) and no user can profit by deviating from his strategy unilaterally. In this paper we focus on a problem of the first kind.

We consider one user who controls dynamically the flow of arriving packets into an infinite buffer. The presence of other users as well as congrstion phenomena is modeled by allowing the service rate to depend on the state of the system, and to 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 case 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. A value iteration algorithm is used to solve this problem in the framework of stochastic zero-sum game. 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 $\mathbf{A} := \{a_1, a_2\}$, $0 \le a_2 \le a_1 \le 1$, 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 $b_1 \ge b_2$. (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 A_t and B_t denote the actions of players 2 and 1 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:

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

(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 (i.e., if $c(0) \geq 0$) 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, γa as a reward related to the acceptance of incoming customer, and θ as a cost per quality of service. Let u be the policy of player 1 and v be the policy of player 2.

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

$$V_\xi(x,u,v) := E^{u,v}\left[\sum_{t=0}^{\infty} \xi^t\, C(X_t, B_t, A_t) \mid X_0 = x\right], \tag{2.2}$$

Define the following problem $(\mathbf{Q}_\xi)$: Find u, v that achieve

$$V_\xi(x) := \sup_{u \in U} \inf_{v \in V} V_\xi(x,u,v), \qquad \forall x \in \mathbf{X}. \tag{2.3}$$

We know from Wessels [13] that there exists a pair of stationary policies (u^*, v^*) that achieves (2.3), and

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

$V_\xi(x)$ is called the ξ-discounted *value* of the game, and the policies (u^*, v^*) are called the optimal policies. For a pair of stationary policies (u,v) let $u_x = \{u_x(1), u_x(2)\} \in M(\mathbf{B})$, where $u_x(i)$ is the probability of choosing b_i; and $v_x = \{v_x(1), v_x(2)\} \in M(\mathbf{A})$, where $v_x(i)$ is the probability of choosing a_i, when (u, v) are used 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.

3 Preliminary Results

Let $\mathcal{K}$ be the set of all real-valued functions on $\mathbf{X}$ that are polynomially bounded. Let the operator $R : \mathbf{X} \times \{1,2\} \times \{1,2\} \times \mathcal{K} \to \mathbb{R}$ be defined as

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

we get:

$$\begin{aligned} R(0, i, j, f) &:= (1 - a_j \bar{b}_i)\, f(0) + a_j \bar{b}_i f(1); && (3.1) \\ R(x, i, j, f) &:= \bar{a}_j b_i f(x-1) + (a_j b_i + \bar{a}_j \bar{b}_i) f(x) + a_j \bar{b}_i f(x+1); \quad x \geq 1. && (3.2) \end{aligned}$$

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

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

and 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

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

The value of the "matrix game D" is defined as $val(D) := \inf_{\alpha \in M(\mathbf{A})} \sup_{\beta \in M(\mathbf{B})} \beta D \alpha$ and is known to satisfy $val(D) = \sup_{\beta \in M(\mathbf{B})} \inf_{\alpha \in M(\mathbf{A})} \beta D \alpha$.

Let $T_\xi : \mathcal{K} \to \mathcal{K}$ be the DP operator associated with $\mathbf{Q}_\xi$:

$$T_\xi f(x) := val\, S(x, f), \tag{3.3}$$

for $x \in \mathbb{N}$.

We shall use the following tools for solving $\mathbf{Q}_\xi$:

Proposition 3.1 *(i) V_ξ satisfies*

$$V_\xi(x) = T_\xi V_\xi(x) \tag{3.4}$$

(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 $T_\xi V_\xi$ in (3.4). Then (u^*, v^*) are optimal for $\mathbf{Q}_\xi$.
(iii) For every $f \in \mathcal{K}$

$$\lim_{n\to\infty} T_\xi^n f = V_\xi, \tag{3.5}$$

Proof. The theorem follows from Wessels [13]). 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

$$\xi |R(x, i, j, \mu^{-1})| \le \alpha \mu^{-1}(x), \quad \forall x, i, j \tag{3.6}$$

Fix some $a \in \mathbf{A}, b \in \mathbf{B}$. Define $\zeta := 1 - \xi^{-1}$ and let

$$r := \frac{\xi^{-1} - ab - \bar{a}\bar{b} + \sqrt{(ab + \bar{a}\bar{b} - \xi^{-1})^2 - 4a\bar{b}\bar{a}b}}{2a\bar{b}}.$$

This choice of r corresponds to a solution of $\xi R(x, i, j, r^x) = r^x$, or equivalently,

$$a\bar{b}r^2 + (ab + \bar{a}\bar{b} - \xi^{-1})r + a\bar{b} = 0. \tag{3.7}$$

We have $\zeta > 0$ and hence $r > 1$. Indeed,

$$r = \frac{\zeta + \bar{a}b + a\bar{b} + \sqrt{(\bar{a}b - a\bar{b})^2 + \zeta^2 + 2\zeta(a\bar{b} + \bar{a}b)}}{2a\bar{b}} \ge \frac{2\zeta + \bar{a}b + a\bar{b} + |\bar{a}b - a\bar{b}|}{2a\bar{b}} > 1 \tag{3.8}$$

Choose some mumber $\alpha > 0$ such that $1 < \alpha r$, and define the function $\mu(x) := (\alpha r)^{-x}$. It follows from (3.7) that $\mu(x)$ satisfies the following:

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

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

which establishes Assumption 2.1 in Wessels [13]. 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 polinomially bounded whereas $\mu^{-1}(x)$ grows exponentially fast (since $\alpha r > 1$). This establishes Assumption 2.2. The Theorem finally follows from from Wessels [13] 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 begin the analysis by listing some properties that a function may possess. Properties **MO** and **IC** bellow are shown in Theorem 4.1 to be sufficient for the optimal policy to have a threshold structure. In order to establish these two properties, we need however to introduce six other properties, which are used to prove that if a function f satisfies **MO** and **IC** then so does $T_\xi f$.

We shall say that $f : \mathbf{X} \times \{1,2\} \times \{1,2\} \to \mathbb{R}$ satisfies assumption:

J1 if $f(x,i,2) - f(x,i,1)$ is monotone nonincreasing in x for $i = 1,2$.

J2 if $f(x,2,j) - f(x,1,j)$ is monotone nondecreasing in x for $i = 1,2$.

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

J4 if $f(x+1,2,1) - f(x+1,2,2) \leq f(x,1,1) - f(x,1,2)$ for $x \in \mathbf{X}$.

J5 if $f(x+1,1,2) - f(x+1,2,2) \leq f(x,1,1) - f(x,2,1)$ for $x \in \mathbf{X}$.

J6 if $f(x+2,i,2) - f(x+1,i,1) \geq f(x+1,i,2) - f(x,i,1)$ for $x \in \mathbf{X}$, $i = 1,2$.

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

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

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

MO if $f(x)$ is monotone nondecreasing in x, i.e. for any $x \in \mathbf{X}$,

$$f(x+1) \geq f(x) \tag{4.2}$$

Let $(u(f), v(f))$ be the stationary policies that achieve the sup and the inf in $valS(x, f(x))$ 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** and **MO**, have the following threshold structure, where randomization occurs in most at one state:

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

$$u_x(f) = \begin{cases} (1,0) & \text{if } x < m_u \\ (q_u, \bar{q}_u) & \text{if } x = m_u \text{ and } m_u = m_v \\ (0,1) & \text{if } x = m_u \text{ and } m_u \neq m_v \\ (0,1) & \text{if } x > m_u \end{cases} \tag{4.3}$$

$$v_x(f) = \begin{cases} (1,0) & \text{if } x < m_v \\ (q_v, \bar{q}_v) & \text{if } x = m_v \text{ and } m_u = m_v \\ (0,1) & \text{if } x = m_v \text{ and } m_u \neq m_v \\ (0,1) & \text{if } x > m_v \end{cases} \tag{4.4}$$

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 $0 \le x$,*

$$F(x) := \zeta_2 h(x+1) + \bar{\zeta}_2 h(x) - \zeta_1 h(x) - \bar{\zeta}_1 h(x-1) \ge 0 \tag{4.5}$$

Proof.

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

■

Lemma 4.2 *Assume that f satisfies* **IC**, **MO**. *Then, $S(\bullet, f)$ satisfies* **J1**- **J6**, *and for any $i, j \in \{1,2\}$, $S(\bullet, i, j, f)$ satisfies* **IC** *and* **MO**.

Proof. We extend $f : \mathbf{X} \to \mathbb{R}$ to $\mathbf{X} \cup \{-1\} \to \mathbb{R}$, and set $f(-1) = f(0)$. With this definition, f satisfies **IC** and **MO** on $\mathbf{X} \cup \{-1\}$, and (3.2) holds for $x = 0$ too. (This extension will save us the need to establish below the boundary conditions).

We begin by **J1**. For $0 \le x$,

$$S(x,i,2,f) - S(x,i,1,f) = (a_2 - a_1)\left[\gamma + \xi\left\{b_i(f(x) - f(x-1)) + \bar{b}_i(f(x+1) - f(x))\right\}\right]$$

Since f satisfies **IC**, the term in curly brackets is monotone nondecreasing, and hence $S(x,i,2,f) - S(x,i,1,f)$ is monotone nonincreasing. This establishes **J1**. The proof of **J2** follows the same arguments.

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

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

since f satisfies **IC**. This establishes **J3**.

Next we check **J4**. For $0 \leq x$,

$$\begin{aligned}
&S(x+1,2,1,f) - S(x+1,2,2,f) - [S(x,1,1,f) - S(x,1,2,f)] \\
&= \; -\xi(a_1 - a_2)\Big(b_2[f(x+1) - f(x)] + \bar{b}_2[f(x+2) - f(x+1)] \\
&\qquad - [b_1[f(x) - f(x-1)] + \bar{b}_1[f(x+1) - f(x)]]\Big) \\
&\leq \; 0
\end{aligned}$$

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

Next we check **J5**. For $0 \leq x$,

$$\begin{aligned}
&S(x+1,1,2,f) - S(x+1,2,2,f) - [S(x,1,1,f) - S(x,2,1,f)] \\
&= \; -\xi(b_1 - b_2)\Big(\bar{a}_2[f(x+1) - f(x)] + a_2[f(x+2) - f(x+1)] \\
&\qquad - [\bar{a}_1[f(x) - f(x-1)] + a_1[f(x+1) - f(x)]]\Big) \\
&\leq \; 0
\end{aligned}$$

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

Next we check **J6**. For $0 \leq x$,

$$\begin{aligned}
&S(x+2,i,2,f) - S(x+1,i,1,f) - [S(x+1,i,2,f) - S(x,i,1,f)] \\
&= \; c(x+2) - c(x+1) - c(x+1) + c(x) + \\
&\quad \xi\Bigg\{ b_i\left[a_2 f(x+2) + \bar{a}_2 f(x+1) - a_1 f(x+1) - \bar{a}_1 f(x)\right] \\
&\quad + \bar{b}_i\left[a_2 f(x+3) + \bar{a}_2 f(x+2) - a_1 f(x+2) - \bar{a}_1 f(x+1)\right] \\
&\quad - b_i\left[a_2 f(x+1) + \bar{a}_2 f(x) - a_1 f(x) - \bar{a}_1 f(x-1)\right] \\
&\quad - \bar{b}_i\left[a_2 f(x+2) + \bar{a}_2 f(x+1) - a_1 f(x+1) - \bar{a}_1 f(x)\right]\Bigg\} \\
&\geq \; \xi\Bigg\{ b_i\left[a_2(f(x+2) - f(x+1)) + \bar{a}_2(f(x+1) - f(x))\right.
\end{aligned}$$

$$-\ a_1(f(x+1)-f(x))-\bar{a}_1(f(x)-f(x-1))]\Bigg\}$$

$$+\xi\Bigg\{\bar{b}_i\left[a_2(f(x+3)-f(x+2))+\bar{a}_2(f(x+2)-f(x+1))\right.$$

$$\left.-a_1(f(x+2)-f(x+1))-\bar{a}_1(f(x+1)-f(x))\right]\Bigg\}$$

$$\geq 0$$

which follows by applying twice Lemma 4.1 with $\zeta_i = a_i$; 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.

The claim that $S(\bullet, i, j, f)$ satisfies **IC** and **MO** is straight forward. ∎

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

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

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

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:

$$l_2^J \leq l_1^J \leq l_2^J + 1 \tag{4.6}$$

$$l_1^I \leq l_2^I \leq l_1^I + 1 \tag{4.7}$$

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 (4.6). 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 (4.7). Since $S(\bullet, f)$ satisfies **J4**, by Lemma 4.2, it follows that

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

and hence the right inequality in (4.6). Finally, since $S(\bullet, f)$ satisfies **J5**, by Lemma 4.2, it follows that

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

and hence the right inequality in (4.7). 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 a_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 a_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 $valS(x,f) = S(x,1,1jf)$.

we consider the following three cases.

Case 1: $l_1^I < l_2^J$. Then

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

Hence, $m_v = l_2^J > m_u = l_1^I$.

Case 2: $l_2^I < l_1^J$. Then

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

Hence, $m_v = l_1^J < m_u = l_2^I$.

Case 3: $l_2^I \ge l_1^J$, and $l_1^I \ge l_2^J$. Then it follows from (4.6) and (4.7) that $m_u = m_v = l_1^I$, $q_u, q_v \in [0,1]$, and

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

where $u_{l_1^I} = \{q_u, \bar{q}_u\}$ and $v_{l_1^I} = \{q_v, \bar{q}_v\}$. 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**, **MO**. *Then $T_\xi f$ satisfies* **IC** *and* **MO**.

Proof. It follows from Wessels [13] Section 4 that $T_\xi f \in \mathcal{K}$. Denote:

$I1 = \{x : x < \min\{m_u, m_v\} - 2\}$,

$I2 = \{x : \min\{m_u, m_v\} \le x < \max\{m_u, m_v\} - 2\}$,

$I3 = \{x : x \ge \max\{m_u, m_v\}\}$.

It follows from Lemma 4.3 that (4.1) holds for $x \in I_1 \cup I_2 \cup I_3$. In order to check in the remaining states, we distinguish between the 3 cases described in the proof of Theorem 4.1.

Denote

$$F(x) = valS(x+2,f) - valS(x+1,f) - [valS(x+1,f) - valS(x,f)],$$

For either

(i) case 1 and $x = \min\{m_u, m_v\} - 2$, (i.e. $x = l_1^I - 2$), or

(ii) case 2, $x = \max\{m_u, m_v\} - 2$, (i.e. $x = l_2^I - 2$), and $m_v = l_1^J < l_2^I - 1 = m_u$ we have

$$F(x) = S(x+2,2,j,f) - S(x+1,1,j,f) - [S(x+1,1,j,f) - S(x,1,j,f)]$$

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

where the last inequality is due to Lemma 4.3, $j=1$ for (i) and $j=2$ for (ii). In case 2, if $x=\max\{m_u,m_v\}-2=l_2^I-2$, and $m_v=l_1^J=l_2^I-1=m_u-1$ then we have

$$
\begin{aligned}
F(x) &= S(x+2,2,2,f)-S(x+1,1,2,f)-[S(x+1,1,2,f)-S(x,1,1,f)] \\
&\geq S(x+2,1,2,f)-S(x+1,1,1,f)-[S(x+1,1,2,f)-S(x,1,1,f)] \\
&\geq 0
\end{aligned}
$$

where the last inequality follows since $S(x,f)$ satisfies **J6**.

For either

(i) case 1, $x=\min\{m_u,m_v\}-1$, (i.e. $x=l_1^I-1$), and $m_v=l_2^J>l_1^I-1=m_u-1$, or

(ii) case 2, $x=\max\{m_u,m_v\}-1$, (i.e. $x=l_2^I-1$), we have

$$
\begin{aligned}
F(x) &= S(x+2,2,j,f)-S(x+1,2,j,f)-[S(x+1,2,j,f)-S(x,1,j,f)] \\
&\geq S(x+2,2,j,f)-S(x+1,2,j,f)-[S(x+1,2,j,f)-S(x,2,j,f)] \\
&\geq 0
\end{aligned}
$$

where the last inequality since by Lemma 4.3, $S(x,f)$ satisfies **J6**; $j=1$ for (i) and $j=2$ for (ii). In case 1, if $x=\min\{m_u,m_v\}-1=l_1^I-1$, and $m_v=l_2^J=l_1^I-1=m_u-1$ then we have

$$
\begin{aligned}
F(x) &= S(x+2,2,2,f)-S(x+1,2,1,f)-[S(x+1,2,1,f)-S(x,1,1,f)] \\
&\geq S(x+2,1,2,f)-S(x+1,1,1,f)-[S(x+1,1,2,f)-S(x,1,1,f)] \\
&\geq 0
\end{aligned}
$$

where the last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **J6**.

For either

(i) case 2 and $x=\min\{m_u,m_v\}-2$, (i.e. $x=l_1^J-2$), or

(ii) case 1, $x=\max\{m_u,m_v\}-2$, (i.e. $x=l_2^J-2$), and $m_v=l_2^J>l_1^I+1=m_u+1$ we have

$$
\begin{aligned}
F(x) &= S(x+2,i,2,f)-S(x+1,i,1,f)-[S(x+1,i,1,f)-S(x,i,1,f)] \\
&\geq S(x+2,i,2,f)-S(x+1,i,1,f)-[S(x+1,i,2,f)-S(x,i,1,f)] \\
&\geq 0
\end{aligned}
$$

where the last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **J6**, $i=2$ for (i) and $i=1$ for (ii). (For case 1, if $x=\max\{m_u,m_v\}-2=l_2^J-2$, and $m_v=l_2^J=l_1^I+1=m_u+1$ we have the situation described above (4.8)).

For either
(i) case 2, $x=\min\{m_u,m_v\}-1$, (i.e. $x=l_1^J-1$), and $m_v=l_2^I>l_1^J+1=m_u+1$,
(ii) case 1, $x=\max\{m_u,m_v\}-1$, (i.e. $x=l_2^J-1$), we have

$$\begin{aligned}F(x) &= S(x+2,i,2,f)-S(x+1,i,1,f)-[S(x+1,i,1,f)-S(x,i,1,f)]\\ &\geq S(x+2,i,2,f)-S(x+1,i,1,f)-[S(x+1,i,2,f)-S(x,i,1,f)]\\ &\geq 0\end{aligned}$$

where the last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **J6**, $i=2$ for (i) and $i=1$ for (ii). (For case 2, if $x=\min\{m_u,m_v\}-1=l_1^J-1$, and $m_v=l_2^I=l_1^J+1=m_u+1$ then we are back in the situation described above (4.8)).

It remains to check Case 3. Denote $\delta_1=(1,0)$, $\delta_2=(0,1)$. For $x=m_u-2=m_v-2=l_1^I-2$ we have

$$\begin{aligned}F(x) &= u_{x+2}S(x+2,f)v_{x+2}-S(x+1,1,1,f)-[S(x+1,1,1,f)-S(x,1,1,f)]\\ &\geq \delta_1 S(x+2,f)v_{x+2}-S(x+1,1,1,f)-[S(x+1,1,1,f)-S(x,1,1,f)]\\ &= q_v\{S(x+2,1,1,f)-S(x+1,1,1,f)-[S(x+1,1,1,f)-S(x,1,1,f)]\}\\ &\quad +\bar{q}_v\{S(x+2,1,2,f)-S(x+1,1,1,f)-[S(x+1,1,1,f)-S(x,1,1,f)]\}\\ &\geq \bar{q}_v\{S(x+2,1,2,f)-S(x+1,1,1,f)-[S(x+1,1,2,f)-S(x,1,1,f)]\}\\ &\geq 0\end{aligned}$$

where the last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **J6**,

For $x=m_u-1=m_v-1=l_1^I-1$ we have

$$\begin{aligned}F(x) &= S(x+2,2,2,f)-2u_{x+1}S(x+1,f)v_{x+1}+S(x,1,1,f)\\ &\geq S(x+2,2,2,f)-u_{x+1}S(x+1,f)\delta_1-u_{x+1}S(x+1,f)\delta_2+S(x,1,1,f)\\ &\geq q_u\{S(x+2,2,2,f)-S(x+1,1,1,f)-S(x+1,1,2,f)+S(x,1,1,f)\}\end{aligned}$$

$$
\begin{aligned}
&+\bar{q}_u\{S(x+2,2,2,f)-S(x+1,2,1,f)-S(x+1,2,2,f)+S(x,1,1,f)\} \\
\geq\ & q_u\{S(x+2,1,2,f)-S(x+1,1,1,f)-S(x+1,1,2,f)+S(x,1,1,f)\} \\
&+\bar{q}_u\{S(x+2,2,2,f)-S(x+1,2,1,f)-S(x+1,2,2,f)+S(x,2,1,f)\} \\
\geq\ & 0
\end{aligned}
$$

The last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **J6**. This establishes that $S(\bullet,f)$ satisfies **IC**.

In order to prove that $valS(\bullet,f)$ satisfies **MO**, it suffices to show that $valS(1,f) \geq valS(0,f)$, since we already established that $S(\bullet,f)$ satisfies **IC**. Indeed,

$$
\begin{aligned}
valS(1,f)-valS(0,f) &= S(1,u_1,v_1,f)-S(0,u_0,v_0,f) \\
&\geq\ S(1,u_0,v_1,f)-S(0,u_0,v_1,f) \geq 0
\end{aligned}
$$

The last inequality follows since, by Lemma 4.3, $S(x,f)$ satisfies **MO**. ∎

We are now ready to prove the main result.

Theorem 4.2 *The optimal value V_ξ satisfies* **MO** *and* **IC**, *and the optimal policies (u^*,v^*) have a threshold structure described by Theorem 4.1.*

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

Acknowledgement: I wish to thank Jerzy Filar for many fruitful discussions and for introducing me into stochastic game.

References

[1] C. Douligeris and R. Mazumdar, "An approach to Flow Control in an Integrated Environment", *CTR Technical Report*, CUCTR-TR-50, Columbia University, 1987.

[2] 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.

[3] C. Douligeris and R. Mazumdar, "More on Pareto Optimal Flow Control," *Proc. of the 26th Allerton Conference* Univ. of Illinois, Urbana, Oct. 1987.

[4] 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.

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

[6] 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.

[7] 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.

[8] M. Marchand, "Prriority Pricing", *Maanagement Science*, **20**, No. 7, pp. 1131-11400, 1974.

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

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

[11] R. Mazumdar, L. Mason and C. Douligeris, "Fairness in Network Optimal Flow Control: Optimality of Product Forms", preprint.

[12] 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.

[13] 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.