Math. Scientist **31**, 35–41 (2006)
Printed in England

A PEST IMMIGRATION PROCESS CONTROLLED BY AN INTERMITTENT PREDATOR

E. G. KYRIAKIDIS * AND
T. D. DIMITRAKOS,** *University of the Aegean*

Abstract

An infinite-state Markov decision model is considered for the control of a simple immigration process, which represents a pest population, by an intermittent predator. It is assumed that the predator may leave the habitat before capturing all the pests. The cost rate caused by the pests is an increasing function of their population size, while the cost rate of the controlling action is constant. A sequence of suitable finite-state Markov decision models is constructed such that the optimal average-cost policies in the sequence converge to the optimal average-cost policy in the original model. There is strong numerical evidence that the optimal policy introduces the predator if and only if the pest population is greater than or equal to some critical size.

Keywords: Markov decision process; simple immigration process; predator; minimum average cost

2000 Mathematics Subject Classification: Primary 90C40
Secondary 60J27

1. Introduction

The Markov decision model is a suitable mathematical model for various problems concerned with the optimal control of a stochastic process. For the relevant theory of this optimization model and various applications, we refer to Chapters 6 and 7 of Ross (1992) and to Chapter 3 of Tijms (1994). We give a description of this model below.

Suppose that a process is observed at time points $t = 0, 1, \dots$ and is classified into one of a possible number of states. The set of possible states is assumed to be finite or denumerable and is denoted by S. If the state of the process at time t is $i \in S$ and the action a is chosen, then regardless of the past history of the process, two things occur: (i) an immediate cost $C(i, a)$ is incurred, and (ii) at time $t + 1$ the process will be in state j with probability $p_{ij}(a)$ A policy is a rule for choosing actions. An important set of policies is the set of stationary policies, where a policy is said to be stationary if the action chosen at time t depends only on the state of the process at time t. A stationary policy f can be represented by a sequence f_i, $i \in S$ where f_i is the selected action, whenever the process is found to be in state i.

The long-run expected average cost per unit time of a policy π is defined as the limit as $t \to \infty$ of the expected cost incurred in the time interval $[0, t]$ divided by t, given that the policy π is employed. If S is finite, and for each stationary policy the associated Markov chain has no two disjoint closed sets, then there exists a stationary policy that minimizes the long-run

Received 31 January 2005; revision received 2 March 2005.
* Postal address: P. Zervou 13, P. Psychico, 15452, Athens, Greece. Email address: kyriak@fme.aegean.gr
** Postal address: Department of Statistics and Actuarial Science, University of the Aegean, Karlovassi, 83200, Samos, Greece. Email address: dimitheo@aegean.gr

expected average cost per unit time for every initial state. The average cost under the optimal policy does not depend on the initial state. The optimal policy can be found by the policy-iteration, the value-iteration and the linear programming algorithms. For a full description of these algorithms we refer to Chapter 3 of Tijms (1994). We present below the value-iteration algorithm, which in most cases can be programmed very easily.

Value-iteration algorithm

Step 0. Choose $V_0(i)$ with $0 \leq V_0(i) \leq \min_a C(i, a)$, for all $i \in S$. Let $n := 1$.

Step 1. Compute the value function $V_n(i),\ i \in S$, from

$$V_n(i) = \min_a \left[C(i, a) + \sum_{j \in S} p_{ij}(a) V_{n-1}(j) \right],$$

and determine $f^{(n)}$ as the stationary policy whose actions minimize the right-hand side of this equation for all $i \in S$.

Step 2. Compute the numbers

$$m_n = \min_{i \in S}[V_n(i) - V_{n-1}(i)] \quad \text{and} \quad M_n = \max_{i \in S}[V_n(i) - V_{n-1}(i)].$$

The algorithm is stopped with the policy $f^{(n)}$ when $0 \leq M_n - m_n \leq \varepsilon m_n$, where ε is a prespecified tolerance number (for example $\varepsilon = 0.001$). The average cost of $f^{(n)}$ is approximated by $(m_n + M_n)/2$. Otherwise go to Step 3.

Step 3. Set $n := n + 1$ and go to Step 1.

The model that we have described above is called a discrete-time Markov decision model. Many practical problems can be modelled as discrete-time Markov decision models by choosing appropriately the state space and action sets. For the case in which the state space is infinite, the computation of the optimal policy can be achieved by Sennott's approximating sequence method (see Sennott (1997), (1999)). In this method, the infinite state model is approximated by a sequence of finite-state Markov decision models. When the times between consecutive decision epochs are not identical but are exponentially distributed, the corresponding decision model is called a continuous-time Markov decision model. It is possible (see e.g. Sennott (1999, p. 245)) to transform a continuous-time Markov decision model to an equivalent discrete-time Markov decision model so that both models have the same optimal policy. In this paper we shall use the approximating sequence method in an infinite-state continuous-time Markov decision model for pest control, as described in the next section.

2. The model

Consider a pest population which grows stochastically in a habitat according to a simple immigration process with immigration rate $\nu > 0$. It is assumed that the cost of the damage caused by the pests is equal to $c_i,\ i \geq 0$, for each unit of time during which the population size is i. It is natural to assume that the sequence $\{c_i\}$ is nondecreasing and $c_0 = 0$. Furthermore, we suppose that $c_i \to \infty$ as $i \to \infty$ and that $c_i \leq Ai^m,\ i \geq 1$, for some positive integer m and some positive number A.

The pest population may be controlled by some action that introduces a predator in the habitat after some random time that is exponentially distributed. The presence of the predator in the habitat immediately stops the immigration of the pests. As soon as the predator is introduced in the habitat, it captures the pests one at a time with constant rate $\sigma > 0$ until their population size is reduced to zero, or the predator leaves the habitat with rate $\theta > 0$ before capturing all the pests. The unit of time has been chosen in such a way that the rate at which the predator is first introduced or re-introduced in the habitat is equal to 1. Thus, when the controlling action is taken, the length of time until the introduction of the predator is exponentially distributed with unit mean. It is also assumed that $\sigma > \nu\theta$ and that the cost of taking controlling action is $k > 0$ per unit time.

Let i and i' denote the states of the process at which the population size of the pests is i, $i \geq 0$, and the predator is either absent or present in their habitat respectively. A stationary policy f is defined by a sequence $\{f_i\}$, $i \geq 0$, where f_i is the action taken when the process is in state i. It is assumed that $f_i = 1$ when the controlling action, which introduces the predator in the habitat, is being taken and $f_i = 0$ when the controlling action is not taken. If the stationary policy $f \equiv \{f_i\}$, $i \geq 0$, is employed, our assumptions imply that we have a continuous-time Markov chain model for the population growth of the pests with state space $S = \{0, 0', 1, 1', \ldots\}$ and the following transitions in a time interval $(t, t + \delta t)$:

$$\begin{array}{llll} i \to i+1 & \text{with probability} & \nu\delta t + o(\delta t), & i \geq 0, \\ i \to i' & \text{with probability} & f_i\delta t + o(\delta t), & i \geq 0, \\ i' \to (i-1)' & \text{with probability} & \sigma\delta t + o(\delta t), & i \geq 1, \\ i' \to i & \text{with probability} & \theta\delta t + o(\delta t), & i \geq 0. \end{array}$$

Our goal is to find the policy which minimizes the expected long-run average cost per unit time among all stationary policies. The decision epochs include the epochs at which an immigration of a pest occurs, and the epochs at which the predator leaves the habitat. Note that this problem, in the case in which the departure of the predator is possible only after the capture of all pests, has been studied in Kyriakidis (2003).

In many Markov decision problems it can be shown that the optimal policy initiates the controlling action if and only if the state of the process is greater than or equal to a critical level. Such a policy is usually called control-limit policy and the critical level the control-limit. It seems intuitively reasonable that in the present problem the optimal policy belongs to the class of control-limit policies $\{P_n : n = 1, 2, \ldots\}$, where P_n is the stationary policy according to which the controlling action is being taken if and only if the pest population size is equal to or exceeds n. A method that in many cases leads to the proof of the optimality of a control-limit policy is the method of successive approximations (see e.g. Ross (1992, p. 147)). This method first establishes the optimality of a control-limit policy for the corresponding finite-horizon discounted cost problem. Then, this structural property is transferred to the infinite-horizon discounted cost problem, and, finally to the average-cost problem. A different approach to the proof of the optimality of a control-limit policy is a parametric analysis introduced by Federgruen and So (1989). According to this approach, it is first shown that an optimal control-limit policy exists when a parameter (possibly fictitious) takes sufficiently small values. This assertion is then extended inductively from interval to interval of the parameter values. In the Markov decision model studied by Kyriakidis (2003), the Federgruen and So technique was applied by varying a fictitious cost incurred during each unit of time that the process was in state $0'$.

In the present model it seems difficult to prove the optimality of a control-limit policy by using the method of successive approximations or the Federgruen and So approach. However, it is possible to find the optimal policy numerically. This can be achieved by transforming the continuous-time Markov decision model to an equivalent discrete-time Markov decision model, and then applying Sennott's (1997) approximating sequence method to it. This method constructs a sequence of finite-state Markov decision models such that the average costs and the average optimal policies in the sequence converge to the optimal average cost and to an optimal policy in the original model. A set of suitable assumptions guarantees the convergence. Extensive numerical results provide strong evidence that the optimal policy in our model is of control-limit type. Note that in the special case in which $c_i = i,\ i \geq 0$, a condition that guarantees the optimality of P_0 was given by Kyriakidis (2003).

The rest of the paper is organized as follows. In the next section our model is transformed to an equivalent discrete-time Markov decision model, and the optimal average-cost policy is obtained by using the approximating sequence method. A numerical example is presented. In Section 4 we verify the assumptions that guarantee the convergence of the algorithm to the optimal policy.

3. Computation of the optimal policy

We transform the continuous-time Markov decision model that was introduced in the previous section to an equivalent discrete-time Markov decision process with the same state space S. This will be achieved by using the standard uniformization technique (see Sennott (1999, p. 245)). Let τ be a positive number such that $\tau < \min\{(\nu+1)^{-1}, (\sigma+\theta)^{-1}\}$. Let $p_{ij}(a),\ a \in \{0,1\}$, be the probability that the next state of the discrete-time process will be j given that the present state is i and the action $a \in \{0,1\}$ is taken, and let $C(i,a)$ be the corresponding one-step expected cost. These parameters are given by (see Sennott (1999, p. 245))

$$\begin{aligned}
p_{i,i+1}(a) &= \tau\nu, && i \geq 0,\ a \in \{0,1\},\\
p_{ii}(0) &= 1-\tau\nu, && i \geq 0,\\
p_{ii'}(1) &= \tau, && i \geq 0,\\
p_{ii}(1) &= 1-\tau(\nu+1), && i \geq 0,\\
p_{i'i}(0) &= \tau\theta, && i \geq 0,\\
p_{i',(i-1)'}(0) &= \tau\sigma, && i \geq 1,\\
p_{i'i'}(0) &= 1-\tau(\sigma+\theta), && i \geq 1,\\
p_{0'0'}(0) &= 1-\tau\theta, &&\\
C(i,0) = C(i',0) &= c_i, && i \geq 0, \qquad (1)\\
C(i,1) &= c_i + k, && i \geq 0. \qquad (2)
\end{aligned}$$

The discrete-time Markov decision model has the same average cost as the original one under any stationary policy. Thus both models have the same optimal policy. Following Sennott's (1997) approximating sequence method, we consider a sequence of finite state space truncations of the above discrete-time Markov decision model. The state space of the truncated model is defined as $G_N = \{0, 0', 1, 1', \dots, N, N'\}$. The one-step transition probabilities $\widetilde{p}_{ij}(a),\ i,j \in G_N$, $a \in \{0,1\}$, of the truncated model coincide with the one-step transition probabilities $p_{ij}(a)$ of the infinite-state model except for the following ones:

$$\begin{aligned}
\widetilde{p}_{NN'}(0) &= \tau\nu,\\
\widetilde{p}_{NN'}(1) &= \tau(\nu+1).
\end{aligned}$$

TABLE 1: The optimal critical points and minimum average costs.

N	n^*	g_{n^*}
70	0	8.1362
72	7	8.2506
75	6	8.4110
80	5	8.6797
85	4	8.9075
200	4	9.3674
250	4	9.3691
290	4	9.3692
300	4	9.3693
1000	4	9.3693

The one-step expected costs $C(i, a)$, $i \in G_N$, $a \in \{0, 1\}$, coincide with those of the infinite-state model. For $N = 1, 2, \dots,$ the average-cost optimal policy and the minimum average cost of each truncated model can be obtained by applying the value iteration algorithm (see Section 1 of the present paper). From Theorem 2.4 in Sennott (1997), it follows that as $N \to \infty$ the minimum average costs and the average-cost optimal policies in the sequence converge to the minimum average cost and to an optimal policy in the infinite-state model. The assumptions that guarantee the convergence in our model will be stated and verified in the next section. In the implementation of the procedure, the number N is increased until both the change in minimum average cost is sufficiently small and the optimal policy is unchanging. From a great number of examples that we have tested, we conjecture that the optimal policy belongs to the class of control-limit policies $\{P_n : n = 1, 2, \dots\}$. This means that there exists a critical point n^* such that the optimal policy prescribes the controlling action if and only if the pest population size is equal to or exceeds n^*.

As an illustration we give a numerical example. We assume that $\sigma = 8$, $\nu = 3$, $\theta = 2$, $c_i = \sqrt{i}$, $i \geq 0$, and $k = 10$. We choose $\tau = 0.05$ and the number 5×10^{-5} as the tolerance number for the relative difference between the upper and lower bounds for the minimum average cost in the value iteration algorithm. In Table 1, we give for various values of N the optimal critical point n^* and the minimum average cost g_{n^*} for the truncated model with state space G_N.

We observe that n^* does not change and g_{n^*} does not change significantly if N is greater than or equal to 200. Hence we conclude that the optimal policy for the original infinite-state model is the policy P_4 with average cost 9.3693.

4. The convergence of the algorithm

According to Proposition 4.1 in Sennott (1997), in our problem the minimum average costs and the optimal policies of the finite-state Markov decision models converge to the minimum average cost and an optimal policy of the infinite-state model if the following two conditions are satisfied.

(i) There exists a stationary policy d in the infinite-state model and a positive number ε such that d induces an irreducible positive recurrent chain on S with finite average cost $g(d)$ for every initial state and the set $D = \{i \in S \mid \text{there exists } a \text{ such that } C(i, a) \leq g(d) + \varepsilon\}$ is finite.

(ii) We have

$$\lim_{N\to\infty} \pi_i(N) = \pi_i, \quad i \in S \qquad \text{and} \qquad \lim_{N\to\infty} g^N_{f\mid N} = g_f,$$

where

- f is the optimal policy of the infinite-state Markov decision model,
- $f \mid N$ is the stationary policy of the finite-state Markov decision model such that $(f \mid N)(i) = f_i,\ i \in G_N$,
- $\pi_i,\ i \in S$, is the steady-state distribution under the policy f,
- $\pi_i(N),\ i \in G_N$, is the steady-state distribution under the policy $f \mid N$,
- g_f is the average cost under the policy f,
- $g^N_{f\mid N}$ is the average cost under the policy $f \mid N$.

Note that condition (i) implies the existence of an optimal policy f, such that the Markov chain under f is positive recurrent (see Proposition 4.1 in Sennott (1997)).

To prove condition (i) we choose $d = P_0$. The stationary probabilities $\widetilde{\pi}_i$ and $\widetilde{\pi}_{i'},\ i \geq 0$, under the policy P_0 in the infinite-state process satisfy the following balance equations subject to the normalizing condition:

$$\begin{aligned}
\widetilde{\pi}_0 &= [1 - \tau(\nu + 1)]\widetilde{\pi}_0 + \tau\theta\widetilde{\pi}_{0'}, \\
\widetilde{\pi}_{0'} &= (1 - \tau\theta)\widetilde{\pi}_{0'} + \tau\widetilde{\pi}_0 + \tau\sigma\widetilde{\pi}_{1'}, \\
\widetilde{\pi}_i &= [1 - \tau(\nu + 1)]\widetilde{\pi}_i + \tau\nu\widetilde{\pi}_{i-1} + \tau\theta\widetilde{\pi}_{i'}, \qquad i \geq 1, \\
\widetilde{\pi}_{i'} &= [1 - \tau(\sigma + \theta)]\widetilde{\pi}_{i'} + \tau\widetilde{\pi}_i + \tau\sigma\widetilde{\pi}_{(i+1)'}, \qquad i \geq 1,
\end{aligned}$$

$$\sum_{i=0}^{\infty} \widetilde{\pi}_i + \sum_{i=0}^{\infty} \widetilde{\pi}_{i'} = 1.$$

After some manipulations we obtain the following expressions:

$$\begin{aligned}
\widetilde{\pi}_i &= \left[\frac{\nu(\theta + \sigma)}{\sigma(\nu + 1)}\right]^i \frac{\theta(\sigma - \nu\theta)}{\sigma(\theta + 1)(\nu + 1)}, \qquad i \geq 0, \\
\widetilde{\pi}_{i'} &= \left[\frac{\nu(\sigma + \theta)}{\sigma(\nu + 1)}\right]^{i-1} \frac{\nu\theta(\sigma - \nu\theta)}{\sigma^2(\theta + 1)(\nu + 1)}, \qquad i \geq 1, \\
\widetilde{\pi}_{0'} &= \frac{\sigma - \nu\theta}{\sigma(\theta + 1)}.
\end{aligned}$$

Note that the condition $\sigma > \nu\theta$, introduced in Section 2, guarantees the existence of the stationary distribution. Clearly, the process under P_0 is irreducible. Furthermore, it is positive recurrent since it has a stationary distribution (see Grimmett and Stirzaker (1992, Theorem 3, p. 205)). The average cost g_0 under P_0 for every initial state is given by (see Tijms (1994, relation (3.1.3)))

$$g_0 = \sum_{i=1}^{\infty} \widetilde{\pi}_i (c_i + k) + \sum_{i=1}^{\infty} \widetilde{\pi}_{i'} c_i.$$

The above expressions for $\widetilde{\pi}_i$, $\widetilde{\pi}_{i'}$, $i \geq 1$, and the condition $c_i \leq Ai^m$, $i \geq 1$, introduced in Section 2, imply that $g_0 < +\infty$. In view of (1), (2) and the assumption that $c_i \to \infty$, as $i \to \infty$, the set $D = \{i \in S \mid \text{there exists } a \text{ such that } C(i, a) \leq g_0 + \varepsilon\}$ is finite for any value of $\varepsilon > 0$.

To prove condition (ii) we note that the balance equations for the states $0, 1, \ldots, N-1, N, 0', \ldots, (N-1)'$ are the same in S and in G_N. They differ only in state N'. This implies that

$$\pi_i(N) = \frac{\pi_i}{\sum_{j \in G_N} \pi_j}, \qquad i \in G_N.$$

From the above expression we deduce that $\lim_{N\to\infty} \pi_i(N) = \pi_i$, $i \in S$. We also have that

$$\begin{aligned} g^N_{f \mid N} &= \sum_{i \in G_N} \pi_i(N) C(i, (f \mid N)(i)) \\ &= \frac{\sum_{i \in G_N} \pi_i C(i, (f \mid N)(i))}{\sum_{j \in G_N} \pi_j} \\ &\to \sum_{i \in S} \pi_i C(i, f_i) = g_f, \quad \text{as } N \to \infty. \end{aligned}$$

References

FEDERGRUEN, A. AND SO, K. C. (1989). Optimal time to repair a broken server. *Adv. Appl. Prob.* **21,** 376–397.

GRIMMETT, G. R. AND STIRZAKER, D. R. (1992). *Probability and Random Processes*, 2nd edn. Oxford University Press.

KYRIAKIDIS, E. G. (2003). Optimal control of a simple immigration process through the introduction of a predator. *Prob. Eng. Inf. Sci.* **17,** 119–135.

ROSS, S. M. (1992). *Applied Probability Models with Optimization Applications.* Dover, New York.

SENNOTT, L. I. (1997). The computation of average optimal policies in denumerable state Markov decision chains. *Adv. Appl. Prob.* **29,** 114–137.

SENNOTT, L. I. (1999). *Stochastic Dynamic Programming and the Control of Queueing Systems.* John Wiley, New York.

TIJMS, H. C. (1994). *Stochastic Models: An Algorithmic Approach.* John Wiley, New York.