A Markov Decision Process Approach to Optimal Control of a Multi-level Hierarchical Manpower System

Akaninyene U. Udom[1]

A recurrent problem in manpower control is how to attain the desired structural configuration in an optimal way, since it is possible to reach a desired structural configuration using different control inputs. The major aim of this paper is to develop a Markov Decision Process for optimal control of a Multi-level Hierarchical Manpower System (MHMS) by promotion and interdepartmental transfers. This is examined under control by intervention and contraction cost Markov Decision Process.

Keywords: Markov Decision Process, Hierarchical Manpower System, Optimal Control.

JEL Classification: C61

1.0 Introduction

The term Markov decision process (MDP) was introduced by Bellman (1957) for the description of a stochastic process controlled by a sequence of actions under conditions of uncertainty. MDP is foundational bridge between stochastic optimal control on one hand and deterministic optimal control on the other. Collections of results with some emphasis to the theoretical aspects of Markov decision processes are given in Derman (1982) and Ross (1992). The most widely used optimization criteria in a Markov decision process are the minimization of the finite-horizon expected cost, the minimization of the infinite-horizon total expected discounted cost or contraction cost, and the minimization of the long-run expected average cost per unit time.

MDP has been used in various aspects of optimization and in different areas, for example Bassey and Chigbu (2012) used MDP approach for the optimal control of oil spill in marine environment from a system-theoretic point of view using the state variable description of Markovian decision process and operational research formalism. Gonzalez-Hermandez and Villareal (2009) give mild conditions for the existence of optimal solution for a Markov decision problem with average cost and m-constraints of the same kind in

[1] Department of Statistics, University of Nigeria, Nsukka, Enugu State, Nigeria. Email: akaninyene.udom@yahoo.com

Borel actions and states spaces. They also show that there is an optimal policy that is a convex combination of at most m+1 deterministic policies whereas in Kyriadis (2011), MDP is used for optimal control of a simple symmetrical pest immigration-emigration process by the introduction of total catastrophes and it was shown that, a particular control-limit policy is average cost optimal within the class of all stationary policies by verifying that the relative values of this policy are the solution of the corresponding optimality equation.

Optimal control is an aspect of optimization in which the input (control) parameters of a dynamical system is manipulated so as to achieve some desired results either by minimizing cost functional or maximizing reward functional associated with the control trajectory of the system. Notable references on this subject on both deterministic and stochastic dynamical systems are Kushner (1972) and Kushner and Runggaldier (1987).

In manpower control, two aspects of control are well known. These aspects of control are attainability (reachability) and maintainability. Whereas attainability is concerned with the process of moving a manpower system from an initial or any given structural configuration to some desired structural configuration, maintainability is concerned with how to remain on the desired structural configuration once it is reached, Bartholomew et al. (1991). Various techniques have been used in optimal manpower control. For example, in Udom and Uche (2009) time is used as an optimality performance criterion, via the Pontryagin minimum principle, to obtain an optimal recruitment control vector for a manpower system modeled by a stochastic differential equation and it was shown that this recruitment vector minimizes the control time globally. Mouza, (2010) adapts a comparative simple dynamic system (plant) with analytical presentation of stocks and flows and proceeds to the formulation of an optimal manpower control problem aiming to achieve in the most satisfactory way, some pre-assigned manpower targets. The work presented a method of solution of the formulated manpower control problem based on the use of the generalized inverse. Other interesting results can be found in Lee et al. (2001), Rao *et al.* (2010) and Nirmala and Jeeva (2010).

A recurrent problem in manpower control is how to attain the desired structural configuration in an optimal way, since it is possible to reach a desired structural configuration using different control inputs. This problem of optimal manpower control is also an issue in the case of a multi-level manpower system. Therefore, the major aim of this paper is to develop a

Markov Decision Process for optimal control of a Multi-level Hierarchical Manpower System (MHMS) by promotion and interdepartmental transfer. This is examined under control by intervention and contraction cost Markov Decision Process.

2.0 Multi-level Hierarchical Manpower System (MHMS)

A manpower system is any identifiable group of people working for the common goal of an organization. Manpower system is usually made of stocks and flows. Stocks refer to the number in the various categorizations of the system and flows comprise of recruitment, promotion, internal transfer and wastage. This section provides the multi-level manpower planning model, as presented in Guerry and De Feyter (2012). For a G-grade manpower system, the stock vector, $n=(n_1,n_2,...,n_G)$, is the vector showing the number of employees in the various grades. A multilevel hierarchical manpower system (MHMS) is a conglomeration of manpower systems made of different departments or levels each made of g-grades. In this case, the stock vector for department or level *d* is $n^d=(n_1^d,n_2^d,...,n_G^d)$, $1\le d\le D$ which represents the number of employees in department *d* in each of the grades. The overall stock for the MHMS is $\underline{n}^\Sigma=(n_1^\Sigma,n_2^\Sigma,...,n_G^\Sigma)$, where $n_g^\Sigma=\sum_{d=1}^{D} n_g^d$ is the overall stock for grade g. With respect to the department or level d, the internal transitions from grade i to grade j are characterized by promotion probability p_{ij}^d. This promotion probability represents the probability that an employee in level d and grade i at time t is in grade j at time t+1in the same department, and this is assumed to be constant in time. The probability that an employee in level d in grade i at time t has left the department at time t+1 is $1-\sum_{j=1}^{G} p_{ij}^d$. For all departments $1\le d\le D$ the promotion matrix is defined by the $D\times D$ block diagonal matrix $P=\begin{pmatrix} P_1 & & & 0 \\ & P_2 & & \\ & & \ddots & \\ 0 & & & P_D \end{pmatrix}$, where $P_d=(p_{ij}^d)$ defines the promotion matrix with respect to level d.

This is the intra-department transition matrix. Transitions between departments are under the control of management and it is assumed that interdepartmental transitions are made at discrete time periods and that this transition can be vertical or horizontal. The transition matrix that represents the evolution of the overall stock vector at organizational level is denoted by P^{Σ}. The maximum likelihood estimator for transition probabilities under Markov assumptions therefore, is given by the following relation:

$$p_{ij}^{\Sigma} = \frac{\sum_{d=1}^{D} n_i^d p_{ij}^d}{\sum_{d=1}^{D} n_i^d}$$

For members of grade s the transition probability from department i to department j is denoted by t_{ij}^{s}. The transitions from department i to department j are then characterized by the diagonal matrix T_{ij} defined by

$$T_{ij} = \begin{pmatrix} t_{ij}^1 & & & \\ & t_{ij}^2 & & \\ & & \ddots & \\ & & & t_{ij}^G \end{pmatrix}$$

For all departments i and j, $(1 \le i, j \le D)$ the information on transition probabilities is contained in the following $D \times D$ block stochastic matrix with t_{ij} as the (i,j)th block.:

$$T = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1D} \\ T_{21} & T_{22} & \cdots & T_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ T_{D1} & T_{D2} & \cdots & T_{DD} \end{pmatrix}$$

This is the interdepartmental transition matrix.

In a similar way, the vector $R^d = (R_1^d, R_2^d, ..., R_G^d)$ is the recruitment vector with respect to department d. The vector $R = (R^1, R^2, \dots, R^D) \in \mathbb{N}^{G,D}$ has as elements the number of recruitments per department and per grade. Now using these notations the stock in department d and grade g at time (t+1) can be predicted using the following recursive equation

$$n_g^d(t+1) = \sum_{i=1}^{D}\sum_{j=1}^{G} n_j^i(t) p_{jg}^i t_{id}^g + R_g^d(t+1)$$

In Guerry and Feyter (2012), attainability of the stock vectors at departmental level of this kind of system is examined under control by recruitment and interdepartmental transitions, in which operationalization of the concept is expressed in terms of realizable approximations and compromise stock vectors that are solutions of constrained optimization problems. They also presented a multi-level optimization algorithm to determine an optimal recruitment strategy resulting in attainable and acceptable stocks that are a compromise between the proposal from the top and the proposals from the departments. A similar type of system is considered in Ossai and Uche (2009) in the context of structural maintainability of the system by introducing the concept of net effect of transfer which is shown to establish the maintainability condition in recruitment control.

3.0 Definitions and problem formulation

Definition 3.1

Let n^d be the stock vector of the multilevel hierarchical manpower system with state space $\mathbf{N}^d$ and let ∂^d be a decision process with control action space $\mathbf{D}^d$. The process $\left(n^d(t), \partial_t^d\right)_{t=0}^{\infty}$ is a MDP if for $n^d \in \mathbf{N}^d$ and $t = 0,1,2,...$ the following holds

$$P\left\{n^d(t+1) = j \,\middle|\, (n^d(0), \partial_0^d), (n^d(1), \partial_1^d), ..., (n^d(t), \partial_t^d)\right\} = P\left\{n^d(t+1) = j \,\middle|\, (n^d(t), \partial_t^d)\right\}$$

Furthermore, for each $k \in \mathbf{D}^d$ and each $n^d \in \mathbf{N}^d$, let $f^d(n^d, k)$ be a cost or penalty vector function associated with level d of the system and P_k^d a Markov matrix of state transition for level d, then $P\left\{n^d(t+1) = j \,\middle|\, (n^d(t) = i, \partial_t^d = k)\right\} = P_k^d(i, j)$ and the cost or penalty $f^d(i,k)$ is incurred whenever $n_t^d = i$ and $\partial_t^d = k$.

Definition 3.2

A decision rule prescribes the procedure for action selection. A deterministic Markov decision rule is a function $\partial^d : \mathbf{N}^d \rightarrow \mathbf{D}^d$ that specifies control action $\partial^d(n^d) \in \mathbf{D}^d$ when the system has the structural configuration n^d. A policy $\pi = (\partial_0^d, \partial_1^d, \partial_2^d, ...)$ in a set of policies Π is a sequence of decision rules, using current information, past information and /or randomization that specifies which control action to take at each point in time. ∂_t^d denotes the decision rule applied at decision epoch $t = 0,1,2,...$. A policy is said to be stationary if $\partial_t^d = \partial^d$. Hence a stationary policy has the form $\pi = (\partial^d, \partial^d, \partial^d, ...)$.

Performance index or decision criterion: This is a function defined on the set of control strategies that represents the tradeoff the decision maker has to make between present and future cost.

Consider a multilevel manpower system whose structural configuration can be influence by a manpower planner by a suitable choice of the control decision variables of the system, namely: promotion, recruitment, interdepartmental transfer and retirement. The structure of the system is observed at specified decision epochs and information necessary to control the system by moving the system from where it is to a desired structural configuration or in the direction of the desired structural configuration is gathered. As a result of the decision to control the system by promotion and interdepartmental transfer, a cost is incurred and consequently, the structural configuration of the system changes to a new structural configuration according to a probability distribution.

We assume that the immediate cost and the transition probability function depend on the structural configuration of the system and the decision taken. It is also assumed that each decision taken is not fully predictable but can be anticipated to some extent before the next action is taken through the probability distribution and any decision taken have a long term consequence on the system. Control decision made at any decision epoch has an impact on control decision at the next epoch and so forth. Each control strategy generates a sequence of costs and thus measures the system response to the decision made. Let the set of feasible structure and decision pair κ be defined by

$$\kappa = \left\{ \left(n^d, \partial^d \right) : n^d \in \mathbf{N}^d, \partial^d \in \mathbf{D}^d \right\}.$$

The decision problem is to find a control policy that minimizes the performance criterion by relating the set of structures $\mathbf{N}^d$, the set of decisions when the structure is n^d, $\mathbf{D}(n^d)$, the transition probability from $n^d(t)$ at time t to $n^d(t+1)$ at time $t+1$, $P^d(n^d(t), n^d(t+1))$, and the cost function $f^d(n^d,k):\kappa \to \mathbf{R}$. Thus the Markov decision for the multi-level manpower system is completely determined by the tuple $\left\{\mathbf{N}^d, \left(\mathbf{D}^d \middle| n^d \in \mathbf{N}^d\right), P^d, f^d(n^d,k)\right\}$.

In this paper, we shall solve this optimal multilevel manpower control problem by considering contraction cost Markov decision criterion and control by intervention.

4.0 Contracted cost Markov decision process

Let the contraction factor be $\alpha \in (0,1)$ and the optimal contraction cost criterion be defined by $V^d(\pi,n^d) = E^{\pi}_{n^d(0)} \sum_{t=0}^{\infty} \alpha^t f(n^d(t), \partial^d_t(n^d(t)))$,

where $E^{\pi}_{n^d(0)}$ denotes expectation conditioned on the initial structure $n^d(0)$ of the multi-level manpower system and the policy π. Then the corresponding optimal contraction value function is

$$V^{d*}(n^d) = \inf_{\pi} V^d(\pi,n^d).$$

The problem is to find the control policy $\pi^* \in \Pi$ such that

$$V(\pi^*,n^d) = V^{d*}(n^d). \qquad (1)$$

A policy $\pi^* \in \Pi$ satisfying (1) is α-optimal.

Theorem 4.1: Let $\pi \in \Pi$ be an arbitrary policy taking decision at the initial time $t=0$. For $\alpha \in (0,1)$, let $W(m^d) \geq \alpha V(n^d)$. Then for all $n^d, m^d \in \mathbf{N}^d$,

$$V^{d*}(n^d) = \min_{\pi \in \Pi} \left[f^d(n^d,\partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right]$$

Proof :

Let $\pi \in \Pi$ be an arbitrary policy taking decision ∂^d_0 at the initial time $t=0$ with the probability of taking this decision p_{∂^d}, $\partial^d \in \mathbf{D}^d$.

Then $V^d(\pi, n^d) = \sum_{\partial^d} p_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) W(m^d) \right]$

Since it is assumed that $W(m^d) \geq \alpha V^d(m^d)$, we have

$$V^d(\pi, n^d) \geq \sum_{\partial^d} p_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right]$$

$$\geq \min_{\partial^d} \sum_{\partial^d} p_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right]$$

Therefore

$$V^d(\pi, n^d) \geq \sum_{\partial^d} p_{\partial^d} \min_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right]$$

$$= \min_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right] \qquad (2)$$

Due to the fact that $\pi \in \Pi$ is arbitrary, inequality 2 implies that

$$V^{d*}(n^d) \geq \min_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right]$$

Now let $\tilde{\partial}^d \in \mathbf{D}$ be some other decision on the structural configuration, then

$$f^d(n^d, \tilde{\partial}^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \tilde{\partial}^d\right) u(m^d) = \min_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right] \quad (3)$$

Supposed that $\tilde{\partial}^d \in \mathbf{D}$ is taken at $t = 0$ on $n^d(0)$ and the expected transition is to m^d then it follows that there exists $\varepsilon > 0$ and $V_\pi^d(m^d)$ such that $V_\pi^d(m^d) \leq V^{d*}(n^d) + \varepsilon$, hence

$$V_\pi^d(n^d) = f^d(n^d, \tilde{\partial}) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \tilde{\partial}\right) V^d(m^d)$$

$$\leq f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) + \alpha\varepsilon$$

Now since $V^{d*}(n^d) \le V^d(\pi, n^d)$, then it is true that

$$V^{d*}(n^d) \le f^d(n^d, \tilde{\partial}^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \tilde{\partial}^d\right) V^d(m^d) + \alpha\varepsilon$$

Therefore, from equation (3) we have

$$V^{d*}(n^d) \le \min_{\partial^d} \left[f^d(n^d, \partial^d) + \alpha \sum_{m \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right] + \alpha\varepsilon$$

The result follows.

An important operator in the analysis of this problem is the operator T_α defined by

$$T_\alpha V^d(n^d) = \min_{\pi \in \Pi} \left[f^d(n^d, \partial^d) + \alpha \sum_{m^d \in \mathbf{N}^d} P\left(m^d \middle| n^d, \partial^d\right) V^d(m^d) \right] \qquad (4)$$

For a finite structural space $\mathbf{N}^d$ and a bounded cost it can be shown that $T_\alpha V^d(n^d)$ is a contraction map on a Banach space B with contraction factor α, hence the contraction map fixed point theorem guarantees the existence of optimal solution V^{d*} to the problem satisfying the optimality equation

$T_\alpha(V^{d*}) = V^{d*}; V^{d*} \in B$.

Definition 4.1

Let B be a real Banach space, that is, a complete normed linear space, with dual B^* and C be a nonempty closed convex subset of B. A mapping $T: C \to C$ is called a contraction map if and only if $\|T(m) - T(n)\| \le \alpha \|m - n\| \ \forall m, n \in C$.

Definition 4.2

A point $m \in C$ is a fixed point of T if $Tm = m$.

Definition 4.3

The modulus of smoothness of B, for $\alpha = 1$, is the function

$\rho_B(t) = \sup\left\{\tfrac{1}{2}\left(\|m+n\|+\|m-n\|-1\right) : \|m\|\le 1; \|n \le t\|\right\}$.

B is uniformly smooth if and only if $\lim_{t\to 0}\dfrac{\rho_B(t)}{t} = 0$.

Theorem 4.2: The operator T_α by equation (4.4) is a contraction map with contraction factor α

Proof: for any $V, W \in B$

$$T_\alpha V^d(n^d) - T_\alpha W^d(n^d) = \min_{\pi\in\Pi}\left[f^d(n^d,\partial^d) + \alpha \sum_{m^d\in\mathbf{N}^d} P\left(m^d \left| n^d, \partial^d\right.\right) V^d(m^d)\right] - \min_{\pi\in\Pi}\left[f^d(n,\partial^d) + \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \partial^d\right.\right) W^d(m^d)\right]$$

(4.5) $$= \min_{\pi\in\Pi}\left[f^d(n,\partial^d) + \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \partial^d\right.\right) V^d(m^d)\right] - \min_{\pi\in\Pi}\left[f^d(n,\bar\partial^d) + \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial^d\right.\right) V^d(m^d)\right]$$

Where $\bar\partial^d$ is such that

$$f^d(n,\bar\partial^d) + \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial^d\right.\right) v(m^d) = \min_{\pi\in\Pi}\left[f^d(n,\partial^d) + \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \partial^d\right.\right) V^d(m^d)\right].$$

Hence from equation (4.5), we have

$$\begin{aligned} T_\alpha V^d(n^d) - T_\alpha W^d(n^d) &\le \alpha \sum_{m^d\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial\right.\right) V^d(m^d) - \alpha \sum_{m\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial\right.\right) V^d(m^d) \\ &= \alpha \sum_{m^d\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial\right.\right)\left[V^d(m^d) - W^d(m^d)\right] \\ &\le \alpha \sum_{m^d\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial\right.\right) \sup_{m\in\mathbf{N}^d}\left[V^d(m^d) - W^d(m^d)\right] \\ &\le \alpha \sum_{m^d\in\mathbf{N}^d} P\left(m^d \left| n^d, \bar\partial\right.\right) \sup_{m\in\mathbf{N}^d}\left|V^d(m^d) - W^d(m^d)\right| \end{aligned}$$

(4.6) $$\le \alpha \left\|V^d - W^d\right\|$$

From equation (4.6) it is easy to see that

$$\sup_{m\in\mathbf{N}^d}\left(T_\alpha V^d(n^d)-T_\alpha W^d(n^d)\right)\le\alpha\left\|V^d-W^d\right\|$$

$$\Rightarrow\sup_{m\in\mathbf{N}^d}\left|T_\alpha V^d(n^d)-T_\alpha W^d(n^d)\right|\le\alpha\left\|V^d-W^d\right\|$$

$$\Rightarrow\left\|T_\alpha V^d(n^d)-T_\alpha W^d(n^d)\right\|\le\alpha\left\|V^d-W^d\right\|$$

4.1 Optimal control by intervention

Here we consider a special type of optimal control in which the multi-level manpower system is allowed to run uncontrolled until the decision maker (manpower planner) chooses to intervene, moving the system instantaneously to some new point in the structural space, from which it is again left uncontrolled until another intervention is made and so on. For any structure n^d in the structural space $\mathbf{N}^d$, let $K(n^d)\in\mathbf{N}^d$ be the set of attainable structures and let $C(n^d,m^d)$ be the cost of moving the system from n^d to m^d and $L(n^d)$ the rate of accumulation of cost in between interventions.

Assumption 4.2

1. There exist a compact set $M^d\subset\mathbf{N}^d$ and a closed set $Z^d\subset\mathbf{N}^d\times M^d$ such that, for all $n^d\in\mathbf{N}^d$, $K(n^d)=\left\{m^d\in M^d:(n^d,m^d)\in Z^d\right\}$, and if $m^d\in K(n^d)$, then $K(m^d)\subset K(n^d)$.
2. C is a continuous function on Z^d.
3. $C(n^d,m^d)+C(m^d,o^d)\ge C(n^d,o^d)$ for $n^d\in\mathbf{N}^d, m^d\in K(n^d), o^d\in K(m^d)$.
4. $0<C_0\le\min_{m^d\in K(n^d)}C(n^d,m^d)\le C_1<\infty$
5. $L:\mathbf{N}^d\to\left[0,C_2\right]$

Let $\tau_1,\tau_2,\tau_3,\ldots$ be the times of intervention, then the average cost incurred will be of the form

$$J_{n^d} = E_{n^d}\left\{\int_{t_1}^{t_2} e^{-\delta t} L(n_t^d)dt + \sum_i e^{-\delta\tau_i} C(n_{\tau_i}^d, m_{\tau_i}^d)\right\} \qquad \delta > 0$$

Then $J_{n^d}(S_{NI}) = E_{n^d}^{S_{NI}}\left\{\int_{t_1}^{t_2} e^{-\delta t} L(n_t^d)dt\right\}$ is the cost of no intervention, which by assumption (4.2) is bounded and coincides with the value function

$$V_0^d(n^d) = E\left\{\int_0^{\tau} e^{-\delta t} L(n_t^d)dt\right\}. \qquad (5)$$

Let Ω be a set of admissible control strategies and Ω_n the set of admissible control strategies such that $\tau_{n+1} = \infty$ almost surely. Then for $\delta > 0$ and under the conditions of assumption (4.2) the cost of a control strategy $S \in \Omega$ when the system is at $n^d \in \mathbf{N}^d$ is

$$J_{n^d}(S) = E_{n^d}^{S}\left\{\int_{t_1}^{t_2} e^{-\delta t} L(n_t^d)dt + \sum_i e^{-\delta\tau_i} C(n_{\tau_i}^d, m_{\tau_i}^d)\right\}.$$

An intervention control strategy $S^* \in \Omega$ is optimal if $J_{n^d}(S^*) = V^{d*}(n^d)$

where $V^{d*}(n^d) = \min_{S\in\Omega} J_{n^d}(S)$.

Theorem 4.3: Let $V_t^d(n^d)$ be a bounded function. If for any strategy $S \in \Omega$, $J_{n^d}(S) = V^d(n^d)$, then,

$$\lim_{t\to\infty} V_t^d(n^d) = V^{d*}(n^d) = \min_{S\in\Omega} J_{n^d}(S)$$

Proof

The limit $W^d(n^d) = \lim_{t\to\infty} V_t^d(n^d)$ certainly exists, since $V_t^d(n^d)$ is bounded. Since $\Omega_n \subset \Omega$, $V_t^d(n^d) \ge V^d(n^d)$ and hence $W^d(n^d) \ge V^d(n^d)$. To get a reversed inequality, we only need to show that the cost of any strategy $S \in \Omega$ can be approximated by that of some strategy $S_n \in \Omega_n$. For any $S \in \Omega$ the intervention times τ_n satisfy $\tau_n \to \infty$ almost surely: this follows from assumption (4.2) on the existence of minimum of intervention cost C_0. Let S_n be the strategy that follows $S \in \Omega$ up to and including time τ_n and then takes no further interventions. Then

$$J_{n^d}(S) - J_{n^d}(S_n) = E^S_{n^d}\left\{\int_{t_n}^{\infty} e^{-\delta t}\left(L(n_t^d) - L(m^d_{t-\tau_n}(\omega_{n+1}))\right)dt + \sum_{i \geq n+1} e^{-\delta \tau_i} C(n^d_{\tau_i}, m^d_{\tau_i})\right\}$$

So that

$$J_{n^d}(S) - J_{n^d}(S_n) \leq E^S_{n^d}\left\{\int_{t_n}^{\infty} \frac{2\|L\|}{\delta} e^{-\delta \tau_n} + \sum_{i \geq n+1} e^{-\delta \tau_i} C(n^d_{\tau_i}, m^d_{\tau_i})\right\}$$

and right-hand side converges to 0 as $t \to \infty$ by dominated convergence. Thus we have shown that $V^*(n^d) = \min_{S \in \Omega} J_{n^d}(S) = \min_{S \in \cup_n \Omega_n} J_{n^d}(S)$

which implies that $W^d(n^d) \leq V^d_t(n^d)$. This completes the proof.

An analysis of the intervention strategies strongly depends on the operator M defined for $\varphi \in B(\mathbf{N}^d)$ (where $B(\mathbf{N}^d)$ is a bounded measurable function on $\mathbf{N}^d$) by

$$M\varphi(n^d) = V^d(n^d) = \min_m [C(n^d, m^d) + \varphi(m^d)$$

where it is assumed that M maps $B(\mathbf{N}^d)$ into itself.

Assumption (4.2) ensures that $\left|M\varphi(n^d)\right| \leq C_1 + \|\varphi\|$ so that $M\varphi(.)$ is bounded.

The properties of this optimal intervention operator is given in the following propositions

Theorem 4.4: Suppose assumption (4.2) are satisfied, then the value function V is the unique fixed point in $B(n^d)$ of the intervention operator G defined by

$$G\varphi(n^d) = \min_\tau E_{n^d}\left(\int_0^\tau e^{-\delta t} L(n^d_t)dt + e^{-\delta\tau} M\varphi(n^d_\tau)\right)$$

Proof: By theorem (4.3) we know that $V_t^d \to V^d$. Where V_0^d is defined by equation (6) and

$$V_{t+1}^d(n^d) = \min_{\tau} E_{n^d}\left\{\int_0^{\tau} e^{-\delta t} L(n_t^d)dt + e^{-\delta\tau} MV_{\tau_n}^d(n_t^d)\right\} \quad (6)$$

Now

$$\begin{aligned} MV^d(n^d) &= \min_{m^d}\left[C(n^d, m^d) + V^d(m^d)\right] \\ &= \min_{m^d}\min_{n^d}\left[C(n^d, m^d) + V^d(m^d)\right] \\ &= \min_{n^d}\min_{m^d}\left[C(n^d, m^d) + V^d(m^d)\right] \\ &= \min_{m^d} MV_t^d(.) \end{aligned}$$

The same argument using the bounded- convergence theorem shows that if φ_t is a decreasing sequence bounded functions and $\varphi_t \to \varphi_\infty$ then $W_t^d \to W_\infty^d$.

Where $W_t(n^d) = \min_{\tau} E_{n^d}\left\{\int_0^{\tau} e^{-\delta t} L(n_t^d)dt + e^{-\delta\tau}\varphi_\tau(n_t^d)\right\}$ for $0 \le t < \infty$. (7)

Taking $\varphi_t = MV_t^d$ and taking limit as $t \to \infty$ on both sides of equation (7) we conclude that V^d satisfies $V^d = GV^d$

Pareto-optimal intervention: Given a MDP and cost function $J_{n^d}(\pi)$ associated with the multi-level manpower system, a control intervention policy $\pi^* \in \Pi$ is Pareto optimal if there is an intervention such that $J_{n^d}(\pi^*) < J_{n^d}(\pi)$.

Theorem 4.5

If $\hat{\pi}(0,T) \in \Pi$ is Pareto-optimal policy for n_0^d then for any $\tau > 0$, $\hat{\pi}(\tau,T)$ is a Pareto-optimal for n_τ^d. Where n_τ^d is the structural configuration of the system at time τ induced by the control intervention $\hat{\pi}(\tau,T)$.

Proof

Consider for n_0^d, $\Pi_i(0)=\left\{\pi \middle| J_j(n_0^d,\hat{\pi}) \le J_j(n_0^d,\pi_0)\right\}$ $j=1,2,...,d \quad i \ne j$. Let $\tau > 0$. We next show that $\hat{\pi}$ minimizes $J_i(n_\tau^d,\pi)$ on the constraint set

$$\tilde{\Pi}_i(0)=\left\{\pi \middle| J_j(n_0^d,\hat{\pi}) \le J_j(n_0^d,\pi_0)\right\} \quad j=1,2,..,d \quad i \ne j$$

To this end, we note that $\hat{\pi}(\tau,T) \in \tilde{\Pi}(0)$, next we show that any element $\pi \in \tilde{\Pi}(\tau)$ can be seen as an element $\pi^* \in \Pi(0)$ restricted to the intervention times (τ,T). That is for all $\pi \in \tilde{\Pi}$ there exist $\pi^* \in \Pi(0)$ such that $\pi^*(\tau,T)=\hat{\pi}$. To that end, let $\pi^*(0,T)$ be the concatenation of $\pi(0,\tau)$ with $\pi(\tau,T)$. Then clearly, $\pi^*(\tau,T)=\hat{\pi}$ is such that $n_\tau^d = n_\tau^{*d}$.

$$J_{n^d}(S)=E_{n^d}^S\left\{\int_0^T e^{-\delta t}L(n_t^d)dt+\sum_i e^{-\delta\tau_i}C(n_{\tau_i}^d,m_{\tau_i}^d)\right\}$$

Furthermore,

$$= E_{n^d}^S\left\{\int_0^\tau e^{-\delta t}L(n_t^d)dt+\int_\tau^T e^{-\delta t}L(n_t^d)dt+\sum_i e^{-\delta\tau_i}C(n_{\tau_i}^d,m_{\tau_i}^d)\right\}$$

$$\le E_{n^d}^{S^*}\left\{\int_0^\tau e^{-\delta t}L(n_t^d)dt+\int_\tau^T e^{-\delta t}L(n_t^d)dt+\sum_i e^{-\delta\tau_i}C(n_{\tau_i}^d,m_{\tau_i}^d)\right\}$$

from dynamic programming principle, it follows directly that $\hat{\pi}$ minimizes $J_i(n_\tau^d,\pi)$ on the constraint set $\tilde{\Pi}_i(0)=\left\{\pi \middle| J_j(n_0^d,\hat{\pi}) \le J_j(n_0^d,\pi_0)\right\}$.

5.0 Illustrative application

The purpose of this section is to illustrate the application of the optimal multi-level hierarchical manpower control model to a manpower data. A 2-level hierarchical manpower system made of 3 grades each is considered. For the system under consideration, as a matter of management policy, promotional control interventions are carried out annually. The following well defined one-step Markovian promotion transition (control) matrices are the control strategies available to the manpower planner for the purpose of either attaining or maintaining a desire structural configuration:

$$P_1^1 = \begin{pmatrix} 0.2 & 0.5 & 0.1 \\ 0.0 & 0.5 & 0.4 \\ 0.0 & 0.0 & 0.7 \end{pmatrix}; \quad P_2^1 = \begin{pmatrix} 0.2 & 0.2 & 0.4 \\ 0.3 & 0.33 & 0.2 \\ 0.0 & 0.5 & 0.2 \end{pmatrix}; \quad P_3^1 = \begin{pmatrix} 0.4 & 0.3 & 0.1 \\ 0.1 & 0.6 & 0.15 \\ 0.1 & 0.05 & 0.8 \end{pmatrix}$$

$$P_1^2 = \begin{pmatrix} 0.6 & 0.15 & 0.1 \\ 0.05 & 0.7 & 0.15 \\ 0.1 & 0.2 & 0.65 \end{pmatrix}; \quad P_2^2 = \begin{pmatrix} 0.0 & 0.3 & 0.4 \\ 0.4 & 0.25 & 0.15 \\ 0.0 & 0.4 & 0.35 \end{pmatrix}; \quad P_3^2 = \begin{pmatrix} 0.1 & 0.6 & 0.3 \\ 0.1 & 0.5 & 0.35 \\ 0.2 & 0.3 & 0.4 \end{pmatrix}$$

The transition matrices above are sub-stochastic matrices as a result of the transitions out of the system (wastage), see also Bartholomew *et al.* (1991).

The cost matrices associated with these promotion control matrices are:

$$V\left(P_1^1\right) = \begin{pmatrix} 300 & 500 & 250 \\ 0 & 1250 & 600 \\ 0 & 0 & 1050 \end{pmatrix}; \quad V\left(P_2^1\right) = \begin{pmatrix} 300 & 200 & 1000 \\ -30 & 825 & 300 \\ 0 & -750 & 800 \end{pmatrix};$$

$$V\left(P_3^1\right) = \begin{pmatrix} 600 & 30 & 250 \\ -100 & 1500 & 225 \\ -250 & -75 & 3200 \end{pmatrix} \quad V\left(P_1^2\right) = \begin{pmatrix} 900 & 150 & 250 \\ -5 & 1750 & 225 \\ -250 & -300 & 2600 \end{pmatrix};$$

$$V\left(P_2^2\right) = \begin{pmatrix} 0 & 300 & 1000 \\ -400 & 626 & 225 \\ 0 & -600 & 1400 \end{pmatrix} \quad V\left(P_3^2\right) = \begin{pmatrix} 150 & 600 & 750 \\ -100 & 1250 & 525 \\ -500 & -450 & 1600 \end{pmatrix}$$

The problem is to obtain the combination of promotion controls from the different levels, such that the cost of control interventions for the system is minimum in the long run (after a large number of intervention epochs). The long run (stationary) distributions of the promotion transitions are obtained by solving the following sets of non-homogeneous equations

$$\xi^d = \xi^d P^d$$

subject to $\sum_{i=1}^{3} \xi_i^d = 1$,

where $\xi^{\mathrm{d}} = \left(\xi_1^d, \xi_2^d, \xi_3^d\right)$.

Table 1 shows the expected cost associated with the different control interventions. The minimum expected costs of control intervention with respect to the levels are asterisked

Levels	Promotion Control Strategies	V_1^d	V_2^d	V_3^d	ξ_1^d	ξ_1^d	ξ_3^d	Expected cost associated with a control strategy V^d	Expected minimum cost for level d V^{d*}	Overall minimum $\sum_d V^{d*}$
	P_1^1	60	875	1000	0	0	1	1000		
$d=1$	P_2^1	51	-62.75	620	$\frac{3}{37}$	$\frac{8}{37}$	$\frac{26}{37}$	426.24^*	**426.24**	
	P_3^1	205	905.25	1338.8	$\frac{11}{50}$	$\frac{33}{50}$	$\frac{6}{50}$	803.22		
										784.78
	P_1^2	514.75	1187.5	1748.8	$\frac{19}{43}$	$\frac{14}{43}$	$\frac{10}{43}$	1020.8		
$d=2$	P_2^2	-160	6.5	713.75	$\frac{16}{119}$	$\frac{40}{119}$	$\frac{9}{17}$	358.54^*	**358.54**	
	P_3^2	-95	1120	1048.75	$\frac{13}{61}$	$\frac{24}{61}$	$\frac{24}{61}$	833.03		

Table 1: The Expected cost of control intervention and the associated optimal control strategy

Table 1 shows the values of the expected cost associated with the promotion control interventions. It can be seen, in the table, that the expected minimum cost of control interventions for the system is **784.78** and the combination of promotion control strategies associated with this optimal value are P_2^1 and P_2^2. The managerial implication therefore is that the control strategies P_2^1 and P_2^2 should be used for optimal promotion control interventions of level 1 and level 2 respectively, since these are the strategies giving minimum expected cost in their respective levels.

6.0 Conclusion

In this paper, we have developed a Markov Decision Process for optimal control of a Multi-level Hierarchical Manpower System (MHMS) by promotion and interdepartmental transfers. This is examined under control by intervention and contraction cost MDP. Theorem on Pareto optimal intervention is presented. An illustrative application on optimal control by annual promotion control intervention of a 2-level hierarchical manpower system made of 3 grades each is also presented. In this way, the problem of optimal manpower attainment is solved by Markov decision formalism.

References

Bartholomew, D.J. Forbes, A.F. and McClean, S.I. (1991). Statistical Techniques for Manpower Planning, John Wiley & Sons.

Bassey, K.J. and Chigbu, P.E. (2012). On optimal control theory in marine oil spill management: A Markovian decision approach, *European Journal of Operational Research,* **217**: 470 – 478.

Bellman, R. (1957). *Dynamic Programming* Princeton University Press, Princeton, NJ, USA.

Derman, D. (1982). *Finite State Markovian Decision Processes*, Academic Press, New York, USA.

Gonzalez-Hermandez, J and Villareal, C.E. (2009). Optimal policies for constrained average-cost Markov decision processes, *Top*, doi 10.1007/s11750-009-0110-7.

Guerry, M-A and T. De Feyter (2012). Optimal recruitment strategies in a multi-level manpower planning model. *Journal of the Operational Research Society* **63:** (7), 931 – 940.

Kushner, H.J. (1972). Necessary conditions for continuous parameter stochastic optimization problems, *SIAM Journal on Optimization and Control* **10**: 550 – 565.

Kushner, H.J. and Runggaldier, W. (1987). Nearly optimal state feedback control for stochastic systems with wideband noise disturbances, *SIAM Journal on Control and Optimization,* **25**: 289 – 315

Kyriadis, E.G. (2011). Explicit solution of the average-cost optimality equation for a pest control problem, Advances in Decision Sciences **2011** doi 10. 1155/2011/617812.

Lee, H.W.J., Cai, X.Q. and Teo, K.L. (2001). An optimal control approach to manpower planning problems, *Mathematical Problems in Engineering* **7**:155 – 175.

Mouza, A.M. (2010). Application of optimal control in manpower planning, *Quality and Quantity*, **44 (**2), 199 – 215.

Nirmala, S. and Jeeva, M. (2010). A dynamic programming approach to optimal manpower recruitment and promotion policies for the two grade system, *African Journal of Mathematics and Computer science* **3**: (12), 297 – 301.

Ossai, E.O. and Uche, P.I. (2009). Maintainability of Departmentalized Manpower Structures in Markov Chain Model Pacific *Journal of Science and Technology* **10**: (2), 295 – 302.

Rao, P.T., Reddy, P.R.S. and Suhasini, A.V.S. (2010). Optimal manpower recruitment by stochastic programming in graded manpower systems, *International Journal of Systems Assurance Engineering and Management* **1**: (2), 179 – 185.

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

Udom, A.U. and Uche, P.I. (2009). The use of time as an optimality performance criterion in manpower control, *Pacific Journal of Science and Technology* **10**: No. (1), 197 – 203