Mathematical Modelling, Vol. 6, pp. 333–338, 1985
Printed in the U.S.A.

0270-0255/85 $3.00 - .00

MULTIOBJECTIVE MARKOV DECISIONS IN URBAN MODELLING

EDWARD BELTRAMI and MICHAEL KATEHAKIS

Department of Applied Mathematics and Statistics
State University of New York at Stony Brook
Stony Brook, New York 11794

SANJA DURINOVIC

Faculty of Economics
Zagreb University
Yugoslavia

Communicated by Edward Beltrami

(*Received January* 1985)

Abstract—A model is developed for deploying emergency services in suburban areas. An area consists of several districts (townships) each of which has its own response units. Each district can employ mutual aid by servicing alarms originating in the others. These decisions are made dynamically. We consider two conflicting measures of performance: the area average response time and the steady-state deterioration in the ability of each district to handle future alarms originating in its own region. We propose a multiobjective Markov Decision formulation and demonstrate how efficient policies can be obtained by solving numerically a specific "small" problem.

Keywords: Markov Decisions, Emergency Services, Urban Modelling, Multiple Objectives.

INTRODUCTION

We develop a model for deploying emergency services in suburban areas. An area typically consists of several districts (townships) each of which has its own response units to service alarms originating within the district. To achieve a better overall level of service the districts can employ mutual aid by servicing alarms originating in other districts even when those districts that receive aid have some units idle. However, by doing so the capability of a district to handle alarms that may occur within it (future alarms) may be reduced. Since districts are primarily responsible for their own areas, a significant deterioration in the immediate level of performance in any district might not be acceptable even if it means a better overall level of performance. For this reason we consider two criteria in our model. The first measures the average response time to current alarms for the whole area. The second is a measure for the deterioration of the capability of each district to handle future alarms in its respective area if mutual aid is employed. For simplicity, we assume that an alarm is always served by a single unit and we suppose that the area consists of two districts only. These assumptions are easy to relax at the expense of increasing the state space. Further it is assumed that if an alarm occurs when all units are busy then an "external" unit responds to this alarm.

The decisions of when to employ mutual aid are made dynamically before the next alarm occurs based on knowledge of the number of busy units and where they have been deployed. This allows the necessary "think" time for a single controller for the whole area.

We provide a multi-objective Markov Decision model for this problem and demonstrate how efficient deterministic policies can be obtained by solving numerically a "small" specific example.

Even though our model is new, Markov Chain and Markov Decision Models have been used in studying questions of emergency systems design and operations. Relevant work in this area includes references [1] to [8].

DESCRIPTION OF THE MODEL

We make the following assumptions: (i) the occurrence of alarms in any district constitutes a stationary Poisson process and the processes corresponding to different districts are independent, (ii) the service time of each unit is exponentially distributed and different service times are independent, (iii) after servicing alarms, units return to the districts of their origin.

A consequence of these assumptions is that the state of the system can be specified by a vector $\mathbf{j} = (j_{11}, j_{12}, j_{21}, j_{22})$ where j_{mn} denotes the number of units from district m busy serving alarms in district n. Furthermore, when a stationary policy is employed, the evolution of the state of the system can be described by a continuous time finite state irreducible Markov Chain. We can consider the following two indices of performance for the two district system.

i) A measure of the steady state average response (travel) time.
ii) A measure of the steady state deterioration of the capability of each district to handle future alarms in its respective region by its own units due to mutual aid.

For the exact definition of the above indices we need the following notation. Let N_i, denote the total number of units in district i $(i = 1, 2)$.

$$S = \{\mathbf{j} = (j_{11}, j_{12}, j_{21}, j_{22}), j_{mn} \geq 0, j_{11} + j_{12} \leq N_1, j_{21} + j_{22} \leq N_2\}$$

S is the state space of the system.

$S_1 = \{\mathbf{j} \in S: j_{11} + j_{12} = N_1 \text{ and } j_{21} + j_{22} = N_2\}$, states in which all units of the area are busy.

$S_0 = S - S_1$ states in which not all units of the area are busy.

λ_i = average rate of demand in district i.

μ_{mn}^{-1} = average service time of a unit from district m serving an alarm in district n.

$e_\pi(\mathbf{j})$ = steady state probability for state $\mathbf{j}$ when a stationary policy π is employed.

$T_{mn}(\mathbf{j})$ = average response time for a unit in district m to reach an alarm in district n when the system is in state $\mathbf{j}$. When $m = n$, these quantities can be obtained according to the "square root formula"[4]. For $m \neq n$, an appropriate extension is straightforward. For simplicity, we assume that $T_{mn}(\mathbf{j}) = T \; \forall \; m,n$ when $\mathbf{j} \in S_1$.

$P_m(\mathbf{j})$ = probability of an alarm occurring in district m when the system is in state $\mathbf{j}$. Notice that

$$P_m(\mathbf{j}) = \frac{\lambda_m}{\lambda_1 + \lambda_2 + j_{11}\mu_{11} + j_{12}\mu_{12} + j_{21}\mu_{21} + j_{22}\mu_{22}} \qquad (m = 1, 2).$$

$P_{mn}(\mathbf{j})$ = probability that when in state $\mathbf{j}$, the next event is a return of a unit of district m after completing service in district n. Then:

$$P_{mn}(\mathbf{j}) = \frac{\mu_{mn} j_{mn}}{\lambda_1 + \lambda_2 + j_{11}\mu_{11} + j_{12}\mu_{12} + j_{21}\mu_{21} + j_{12}\mu_{22}} \qquad \text{if } \mathbf{j} \in S_0.$$

and

$$P_{mn}(\mathbf{j}) = \frac{\mu_{mn} j_{mn}}{j_{11}\mu_{11} + j_{12}\mu_{12} + j_{21}\mu_{21} + j_{22}\mu_{22}} \qquad \text{if } \mathbf{j} \in S_1.$$

When the system is in state $\mathbf{j} \in S_0$, based on the number of free units available at least one of the following actions can be taken.

a_0: Service the next alarm by a unit in the district of the origin of the alarm.

a_1: Service the next alarm regardless of the origin with a unit from district 1.

a_2: Service the next alarm regardless of origin with a unit from district 2.

Finally, let

$q_m(\mathbf{j}, a)$ denote the conditional probability of a future alarm occurring in district m given that in state $\mathbf{j}$ action α was taken and an alarm in district $m' \neq m$ occurred. Note that by "future" we mean that an alarm occurs before the next return of a unit takes place.

It is easy to find these probabilities. For instance,

$$q_1(\mathbf{j}, a_1) = \frac{\lambda_1}{\lambda_1 + \lambda_2 + j_{11}\mu_{11} + (j_{12} + 1)\mu_{12} + j_{21}\mu_{21} + j_{22}\mu_{22}}$$

Notice that transitions from state $\mathbf{j}$ involve a change of exactly one of the variables $j_{mn}(m, n = 1, 2)$ to $j_{mn} + 1$ or $j_{mn} - 1$, and changes to $j_{mn} + 1$ are possible only if $j_{m1} + j_{m2} < N_m(m = 1, 2)$. The transition probabilities $p(\mathbf{j}/\mathbf{i}, a)$ from state $\mathbf{i}$ to state $\mathbf{j}$, given a feasible action a can be readily computed. As an example, consider the state $i_. = (2, 1, 0, 2)$ when $N_1 = 4$ and $N_2 = 2$. The only action feasible is alternative a_1 and the following transitions are possible:

to state	*with transition probability*
(3, 1, 0, 2)	$P_1(\mathbf{j})$
(2, 2, 0, 2)	$P_2(\mathbf{j})$
(1, 1, 0, 2)	$P_{11}(\mathbf{j})$
(2, 0, 0, 2)	$P_{12}(\mathbf{j})$
(2, 1, 0, 1)	$P_{22}(\mathbf{j})$

For each state and action taken we define the following "costs":

i) Expected response times:

$$C^1(\mathbf{j}, a_0) = P_1(\mathbf{j})T_{11}(\mathbf{j}) + P_2(\mathbf{j})T_{22}(j) \qquad \text{if } \mathbf{j} \in S_0,$$

$$C^1(\mathbf{j}, a_1) = P_1(\mathbf{j})T_{11}(\mathbf{j}) + P_2(\mathbf{j})T_{12}(j) \qquad \text{if } \mathbf{j} \in S_0,$$

$$C^1(\mathbf{j}, a_2) = P_1(\mathbf{j})T_{21}(\mathbf{j}) + P_2(\mathbf{j})T_{22}(\mathbf{j}) \qquad \text{if } \mathbf{j} \in S_0,$$

and

$$C^1(\mathbf{j}) = (P_1(\mathbf{j}) + P_2(\mathbf{j}))T \qquad \text{if } \mathbf{j} \in S_1$$

ii) Expected incremental increase in response times for the individual districts to handle future demands by their own means in their respective regions due to mutual aid.

If $\mathbf{j} \in S_0$, define:

$$C^2(\mathbf{j}, a_0) = 0$$

$$C^2(\mathbf{j}, a_1) = \begin{cases} P_2(\mathbf{j})q_1(\mathbf{j}, a_1)[T_{11}(j_{11}, j_{12} + 1, j_{21}, j_{22}) - T_{11}(\mathbf{j})] & \text{if } j_{11} + j_{12} + 1 < N_1 \\ P_2(\mathbf{j})q_1(\mathbf{j}, a_1)[T_{21}(j_{11}, j_{12}, j_{21} + 1, j_{22}) - T_{11}(\mathbf{j})] & \text{if } j_{11} + j_{12} + 1 = N_1 \end{cases}$$

$$C^2(\mathbf{j}, a_2) = \begin{cases} P_1(\mathbf{j})q_2(\mathbf{j}, a_2)[T_{22}(j_{11}, j_{12}, j_{21} + 1, j_{22}) - T_{22}(\mathbf{j})] & \text{if } j_{21} + j_{22} + 1 < N_2 \\ P_1(\mathbf{j})q_2(\mathbf{j}, a_2)[T_{12}(j_{11}, j_{12}, j_{21} + 1, j_{22}) - T_{22}(\mathbf{j}_1)] & \text{if } j_{21} + j_{22} + 1 = N_2 \end{cases}$$

and if $\mathbf{j} \in S_1$, we define $C^2(\mathbf{j}) = 0$.
Notice that $C^2(\mathbf{j}, a)$ represent a measure of "opportunity cost" associated with taking action a when the system is in state $\mathbf{j}$.

COMPUTATION OF EFFICIENT DETERMINISTIC POLICIES AND NUMERICAL RESULTS

When a deterministic policy π is followed, the cost structure defined in the previous section implies the following two measures of performance.

1) Average response time:

$$\phi^1(\pi) = \sum_{\mathbf{j} \in S} C^1(\mathbf{j}, \pi(\mathbf{j}))e_\pi(\mathbf{j}).$$

2) Average opportunity cost due to mutual aid:

$$\phi^2(\pi) = \sum_{\mathbf{j} \in S} C^2(\mathbf{j}, \pi(\mathbf{j}))e_\pi(\mathbf{j}).$$

A deterministic policy assigns actions as a deterministic function of the state[9]. Thus, $\pi(\mathbf{j})$ above denotes the action prescribed by policy π in state $\mathbf{j}$ and by convention if $\mathbf{j} \in S_1$, we set $C^1(\mathbf{j}, \pi(\mathbf{j})) = C^1(\mathbf{j})$ and $C^2(\mathbf{j}, \pi(\mathbf{j})) = 0$.

A policy π^0 is said to be efficient if and only if there is no other policy π with the property that:

$$\phi^k(\pi) \le \phi^k(\pi^0) \qquad k = 1, 2$$

with strict inequality holding for at least some k. Note that under deterministic policies the resulting Markov Chain is irreducible. In [10], [11], it was established that efficient deterministic policies can be obtained from the extreme efficient points of the following multi-objective linear program (MOLP):

$$\min\{\sum_{\mathbf{j}, a} C^1(\mathbf{j}, a)x_{\mathbf{j}, a}, \sum_{\mathbf{j}, a} C^2(\mathbf{j}, a)x_{\mathbf{j}a}\}$$

subject to

$$\sum_a x_{\mathbf{i}a} - \sum_{\mathbf{j}, a} x_{\mathbf{x}a}p(\mathbf{i}/\mathbf{j}, a) = 0 \qquad \mathbf{i} \in S$$

$$\sum_{\mathbf{i}, a} x_{\mathbf{i}a} = 1$$

$$x_{\mathbf{i}a} \ge 0.$$

Now the (MOLP) can be solved by standard methods[12], [13], [14]. We folllow the approach of forming a single objective linear program by multiplying the first objective by a parameter λ and the second by $(1 - \lambda)$ to form a single scalar objective.

We then employ a parametric cost analysis.

For the specific example of two districts, we developed two computer programs. The first one converts the model into the corresponding (MOLP) problem and writes out the objective functions, the constraint matrix and the right hand side. The only input it requires are the values of the parameters N_1, N_2, λ_1, λ_2, μ_{11}, μ_{12}, μ_{21}, μ_{22} and the average response times. This program is very useful because even for a small problem such as $N_1 = 2$ and $N_2 = 2$, there are 36 constraints and 54 variables. Writing these constraints by hand would be a very tedious task indeed.

The second program takes as input the output of the first program. It combines the two objectives into one by multiplying the first by λ and the second by $(1 - \lambda)$, obtains the optimal solution and then determines, through sensitivity analysis, the range of λ for which this solution remains optimal. All optimal solutions (and thus all efficient extreme points of the problem (MOLP) which give all deterministic efficient policies for the model) are obtained for $\lambda \in [0, 1]$. Note that any edge joining two adjacent efficient extreme points of the problem (MOLP) is also efficient, [14]. Points on such an edge give stationary efficient policies but since for our model only deterministic policies are of practical use, these are not determined.

We present a solution for $N_1 = 2$ and $N_2 = 2$. The other parameters are chosen arbitrarily as

$$\lambda_1 = 4.0 \qquad \lambda_2 = 5.0 \qquad \lambda_{11} = 8.0$$

$$\mu_{12} = 7.0 \qquad \mu_{21} = 6.0 \qquad \mu_{22} = 8.0$$

$$t_{11}(0) = t_{12}(0) = t_{21}(0) = t_{22}(0) = 5.65$$

$$t_{11}(1) = 3.04 \qquad t_{11}(2) = 2.90$$

$$t_{12}(1) = 4.20 \qquad t_{12}(2) = 3.92$$

$$t_{21}(1) = 4.85 \qquad t_{21}(2) = 4.68$$

$$t_{22}(1) = 3.60 \qquad t_{22}(2) = 3.20$$

where the response times are obtained by the formula

$$T_{mn}(s) = t_{mn}(N_m - j_{m1} - j_{m2}) \qquad m, n = 1, 2, s \in S.$$

Table 1

States belonging to the set S_0.	Alternatives available	Action specified by the efficient policy R_i R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}
(0, 0, 0, 0)	a_0, a_1, a_2	a_1	a_0	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2
(0, 0, 0, 1)	a_0, a_1, a_2	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2
(0, 0, 1, 0)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2	a_2	a_2
(1, 0, 0, 0)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2	a_2
(0, 1, 0, 0)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2
(1, 0, 0, 1)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2
(0, 1, 0, 1)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2
(1, 0, 1, 0)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2
(0, 1, 1, 0)	a_0, a_1, a_2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2

Note that policy R_1 minimizes the second objective while policy R_{11} minimizes the first one.

For these values of the parameters eleven optimal solutions were obtained. These solutions together with the values of the objectives and the range of λ for which they remain optimal, are given in Table 1. The eleven corresponding efficient deterministic policies for the model are summarized also in Table 1.

REFERENCES

1. E. Beltrami, Markov decisions in urban modelling. *Math. Modelling* **2**, 213–225 (1981).
2. J. Jarvis, Optimal assignments in a Markovian queueing system. *Comput. and Ops. Res.* **8**, 17–23 (1981).
3. M. Katehakis and A. Levine, Allocation of distinguishable servers asymptotic results, *Comput. and Ops. Res.* ~~(to appear, 19~~85).
4. P. Kolesar and E. Blum, Square root laws for the fire engine response distances I: Background and theory, *Mgmt. Sci.* **19**, 1365–1378 (1973).
5. P. Kolesar, W. Walker and J. Hausner, Determining the relation between fire engine travel times and travel distances in New York City. *Oper. Res.* **23**, 614 (1975).
6. R. C. Larson, A hypercube queueing model for facility location and redistricting in urban emergency services, *Comp. and Op. Res.* **1**, 67–95 (1974).
7. A. Swersey, A Markovian decision model for deciding how many units to dispatch, in *Models for Reducing Fire Engine Response Times*. Ph.D. Thesis, Columbia University, New York (1972).
8. W. Walker, J. Chaiken, and E. Ignall (Editors) *Fire Department Deployment Analysis*, Elsevier North Holland, New York (1979).
9. C. Derman, *Finite State Markovian Decision Processes*, Academic Press, New York (1970).
10. S. Durinovic, J. Fillar, M. N. Katehakis, and Y. Lee, *On Multiobjective MDPs with Average Cost Criteria*, Johns Hopkins Univ. Tech. Report (1984).
11. S. Durinovic, *On Multiobjective Markov Decision Processes*. Ph.D. Thesis, S.U.N.Y. at Stony Brook (1983).
12. J. Cohon, *Multiobjective Programming and Planning*, Academic Press, New York (1978).
13. P. Yu and M. Zeleny, The set of all nondominated solutions in linear cases and a multiobjective simplex method, *J. Math. Anal. Appl.* **49**, 430–468 (1975).
14. M. Zeleny, *Linear Multiobjective Programming*. Springer-Verlag, New York (1974).