REGIONAL SCHOOL DISTRICTING
VIA MATHEMATICAL PROGRAMMING

C. De Giorgi, G. Tagliabue

TECHINT - COMPAGNIA TECNICA INTERNAZIONALE - Milano

P. Migliarese, P.C. Palermo

ISTITUTO DI ELETTRONICA, POLITECNICO di Milano

INTRODUCTION

This paper outlines a regional districting model, which concerns the spatial organisation of the Italian High Educational System, according to a national program implemented since november '74.

The main objects of the program, aiming to perform the regional balance and the efficiency of the educational structures, were:

- to determine the location and the sizes of polyvalent school units (in the following, district centres), supplying all the kinds of High Education and satisfying the regional demand;
- to partition the Regions into school districts in order to share the student population upon district centres according to the main goals of the program.

The structure of this problem corresponds to a well known family of mathematical models (location - allocation models): given a number n of areas, each with a known demand of services or goods, and a number m of alternative sites where the supply centres may be built to satisfy these demands, determine the location and the size of the centres and which demand areas are to be served by a given centre, in order to optimize a given objective function. Obviously, the real problem cannot be solved by mathematical programming model, i.e. by a pure quantitative approach, but such a model may be useful developed within a suitable methodological framework, connecting empiric-qualitative analyses and evaluations with automatic computations.

In this sense a multistage and iterative procedure has been proposed in

Fig. 1

previous works /1/; an essential logical scheme is reported in Fig. 1. Here we will remark only that:

- parameters and constraints of the model are determined in the analysis phase;
- the first level zoning takes into account socio-political and "strategic" factors (often qualitative) and assures the feasibility of the model from a computational point of view in the meantime;
- the main task of the model is to assure the economic efficiency of the plan, given dominant political and social options;
- the iteration of the procedure, as in a simulation process, define a pay-off rate between socio-political goals and economic costs;
- the output of the study isn't "the plan", rather a balance sheet of qualitative-quantitative evaluations of some relevant alternatives.

In this framework a capacited plant location model with capacity indivisibility and side constraints has been developed.

In the following the structure of the model and the algorithm are presented, while experiments and computational results are described in other works /1/,/2/.

MODEL FORMULATION

With reference to a given zone the following entities are defined:

- a) - a set $J = \{j / j = 1, \dots, m\}$ where j represents a potential district centre
- a vector $q_j = \{q_{jk} / k = 1, \dots, k_j\}$ where q_{jk} represents the k -th capacity level of centre j ($k = 1$ corresponds to a closed centre)
 - a vector $f_j = \{f_{jk} / k = 1, \dots, k_j\}$, where f_{jk} represents the production cost for centre j of k -th level
 - a location variable vector $y_j = \{y_{jk} / k = 1, \dots, k_j\}$, $y_{jk} = \begin{cases} 0 \\ 1 \end{cases} \begin{matrix} \\ 1 \end{matrix}$, where if $y_{jk} = 1$ centre j is opened of k -th level

$$\sum_{k=1}^{k_j} y_{jk} = 1, \forall j \in J \quad 2)$$

- b) - a matrix $\{h_{js}\} [m \times m]$, $h_{js} = \begin{cases} 0 \\ 1 \end{cases}$, representing the compatibility among potential district centres, i.e. if $h_{js} = 1$ and s and j are mutually exclusive

$$\sum_{s=1}^m h_{js} (1 - y_{s1}) \leq 1, \forall j \in J \quad 3)$$

- c) - two parameters NMIN and NMAX representing lower and upper bounds for the zone district number

$$NMIN \leq \sum_j (1 - y_{j1}) \leq NMAX \quad 4)$$

- d) - a set $I = \{i/i = 1, \dots, n\}$, where i represents elementary areas, with demand r_i .

The problem is feasible if the following data relation is verified:

$$\sum_{i=1}^n r_i \leq \sum_{j=1}^m q_{jk} y_{jk}$$

- e) - an allocation variable matrix $\{x_{ij}\} [n \times m]$, where x_{ij} $0 \leq x_{ij} \leq r_i$ 5), represents the ratio of demand r_i allocated to centre j . Obviously the allocation must not exceed the centre capacity, i.e.

$$\sum_{i=1}^n x_{ij} \leq \sum_{k=1}^{k_j} q_{jk} y_{jk}, \forall j \in J \quad 6)$$

The demand of each centre must be satisfied:

$$\sum_{j=1}^m x_{ij} = r_i \quad \forall i \in I \quad 7)$$

- f) - a transportation cost matrix $\{c_{ij}\} [n \times m]$
 $c_{ij} = \infty$ if i cannot be allocated to j

g) - the objective function

$$F(x, y) = (1 - \beta) \sum_{j=1}^m \sum_{k=1}^{k_j} f_{jk} y_{jk} + \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij}$$

i.e. the weighted sum between amortized production and transportation cost.

Our problem corresponds to:

$$\min_{x, y} F(x, y) \text{ subject to constraints } 1, \dots, 7$$

Parameter β in the objective function permits to experiment different pay-off rates between public and private costs of school service.

THE ALGORITHM

A two level tree-search algorithm has been developed. First of all a tree (main tree) is scanned, where each node represents a solution of the location problem:

$$\text{determine } y_{jk} = \begin{cases} 1 \\ 0 \end{cases}, \quad \forall j \in J, \quad k = \begin{cases} 1 \\ k_j \end{cases}$$

In the initial node $y_{jk_j} = 1, \forall i \in J$; the centres are progressively "closed" ($y_{j1} = 1, j$ given) according to some preference ordering.

Let J_0^P be, at a given node P, the set of "opened" centres

$$(y_{jk_j} = 1, j \in J_0^P)$$

The node P is feasible if:

- the constraints 2,3,4 (side constraints) are satisfied and

$$-\sum_{i \in I} r_i \leq \sum_{j \in J_0^P} q_{jk_j} \quad 8)$$

In this case the transportation subroutine is applied in order to determine a feasible solution of the model. Otherwise the node P is fathomed and a backtrack process occurs. A secondary tree is generated in corrispondence with each feasible node of the main tree, where the

capacity levels of the "opened" centres are tried to be reduced.

The node PS, corresponding to node P of the main tree, represents a solution of the location problem:

$$\text{determine } y_{jk} = \begin{cases} 1 \\ 0 \end{cases}, \quad j \in J_0^P, \quad k = 2, \dots, k_j$$

The node PS may be fathomized if:

- a feasibility check analogous to 8) is not satisfied or
- an usual bounding condition occurs.

Otherwise the transportation subroutine is applied.

The lower bound, at the node PS of the secondary tree deriving from node P of the main tree, is the sum of two components:

- a transportation cost defined as

$$\min_{x_{ij}} \sum_i \sum_{j \in J_0^P} c_{ij} x_{ij} \quad \text{with constraints}$$

$$\sum_i x_{ij} \leq q_{jk_j} \quad \forall j \in J_0^P$$

$$\sum_{i \in J_0^P} x_{ij} = r_i \quad \forall i \in I$$

$$x_{ij} \geq 0 \quad \forall i \in I, j \in J_0^P$$

- a fixed cost defined as

$$\sum_{j \in J^{PS}} f_{jk(j)} + CF, \text{ where}$$

$J^{PS} = \{j \mid J_0^P / y_{jk(j)} = 1\}$ i.e. the set of opened centres whose capacity level $k(j)$ is fixed on the secondary tree.

$$CF = \min_{z_{jk}} \sum_{k=1}^{k_j} \sum_{j \in \{J_0^P - J^{PS}\}} f_{jk} z_{jk} \quad \text{with constraints}$$

$$z_{jk} = \begin{cases} 0 \\ 1 \end{cases}$$

$$\text{and } \sum_{i=1}^n r_i - \sum_{j \in J^{PS}} q_{jk}(j) \sum_{k=1}^j \sum_{j \in \{J_0^P - J^{PS}\}} q_{jk} z_{jk}$$

The fixed cost is the sum of two terms. The former is the cost of the capacity levels of the nodes backward and that of the node PS, namely of the nodes whose capacities are already fixed. The latter, here denoted as CF, is the minimum cost referring to capacity levels of the nodes forward that satisfy the demand.

CF is computed via Parametric Dynamic Programming because it must be evaluated several times corresponding to different levels of residual demand.

DETAILED ALGORITHM

Define:

$$R = \sum_{i=1}^n r_i$$

$$y = \{y_1, y_2, \dots, y_m\}$$

$$q(y) = \sum_{j=1}^m q_j^T y_j$$

$$h_i, H_i \text{ fathomization indexes}$$

INITIALIZATION

- 1) $\bar{z} = \infty$; $h_j = -1$, $y_{jk_j} = 1 \forall j$
- 2) if $q(y) \geq R$ go to 3) else ERROR go to 9)
- 3) if y satisfies side constraints go to 4) else go to 8)
- 4) $z = F(x, y)$ (x generated via transportation subroutine)
- 5) if $z < \bar{z}$ go to 6) else go to 7)
- 6) $\bar{z} = z$
- 7) CALL TREE 2
- 8) CALL TREE 1

9) STOP

TREE 1

- 1) $j = m; y_{j1} = 1 \forall j$
- 2) if $h_j < 0$ go to 3) else go to 11)
- 3) $h_j = 0$
- 4) if y satisfies side constraints go to 5) else go to 10)
- 5) if $q(y) \geq R$ go to 6) else go to 9)
- 6) $z = F(x, y)$
- 7) if $z < \bar{z}$ go to 8) else go to 9)
- 8) $\bar{z} = z$
- 9) CALL TREE 2
- 10) if $j \neq m$ go to 1) (branch) else go to 11) (backtrack)
- 11) $h_j = -1; y_{jk_j} = 1; j = j-1; y_{j1} = 1$
- 12) if $j = 0$ go to 13) else go to 2)
- 13) RETURN

TREE 2

- 1) $H_j = -1 \forall j \mid y_{jk_j} = 1$
- 2) $j = m^*$, where m^* is the number of opened centres
- 3) if $H_j < 0$ go to 4) else go to 13) (backtrack)
- 4) $H_j = k_j - \bar{k} - 1$, where $\bar{k} = \min_k q_{jk} \mid q_{jk} \geq R - \sum_{s \neq j} q_s^T y_s$
- 5) if $H_j > 0$ go to 6) else go to 13) (backtrack)
- 6) $F_{jk}^* = \min_k LB_{jk}$, $\bar{k} \leq k \leq k_j$, where LB_{jk} is a lower bound on $F(x, y)$ at node $(\{y - y_j\}, y_{jk})$
- 7) if $F_{jk}^* \leq \bar{z}$ go to 8) else go to 13) (backtrack)
- 8) $LB_{jk}^* = \infty; H_j = H_j - 1$
- 9) $z = F(x, y)$
- 10) if $z < \bar{z}$ go to 11) else go to 12)


```

11)  $\bar{z} = z$ 
12) if  $j \neq m^*$  go to 2) (branch) else go to 5)
13)  $H_j = -1$ ;  $j = 1$ 
14) if  $j > 0$  go to 15) else go to 17)
15) if  $H_j = 0$  go to 13) else go to 16)
16) if  $H_j < 0$  go to 4) else go to 6)
17) RETURN

```

This algorithm has been coded in FORTRAN V and tested on UNIVAC 1108, as in reference /1/.

REFERENCES

- /1/ P. Maggiolini, P.C. Palermo, G. Tagliabue, P.P. Vecchi
 "Regional Planning of Educational Systems: Methodology and Mathematical Models" AIRO Symposium, Milan, 1975 (in italian)
- /2/ C. De Giorgi, P.C. Palermo, G. Tagliabue, M. Valdata
 "School Districting Model" AIRO Symposium, Milan 1975 (in italian)
- /3/ C. Revelle, D. Marks, J.C. Liebman
 "An Analysis of Private and Public Sector Location Models"
 Manag. Sci. Vol. 16, N. 11, 692-707 (1970)
- /4/ G. Sà
 "Branch - and - Bound and Approximate Solution to the Capacited Plant-Location Problem" O.R., Vol. 17 1005-1015 (1969)
- /5/ R.J. O'Brien
 "Model for Planning the Location and Size of Urban School" Socio. Econ. Plan. Sci., Vol. 2, 141-153 (1969)