Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 108 (2014) 96 – 105

Procedia
Social and Behavioral Sciences

AIRO Winter 2013

A location model for the reorganization of a school system: the Italian case study.

Giuseppe Bruno[a], Andrea Genovese[b], Carmela Piccolo[a], Claudio Sterle[a]

[a]*University of Naples "Federico II", Piazzale Tecchio 80, 80125, Naples, Italy*
[a]*University of Sheffield, Management School, Conduit Street, Sheffield, UK*

Abstract

This work analyses a problem concerning the reorganisation of the school system located in a given region. In particular, the problem has been tackled considering the requirements indicated for the Italian case related to the reorganisation of pre-primary, primary and secondary schools in integrated institutions managed in a centralised way. In order to solve the problem, two versions of a location model are proposed. The application to a real case study shows how the model can be used as a viable decision support system.

*Keywords:*location; clustering; schools; reorganization.

1. Introduction

In the current economic climate, characterized by growing cuts to public expenditure, public services (e.g., healthcare, education, policing) have undergone significant transformations (Sancton, 2000). Such deep changes have been generally oriented to reduce administrative, managerial and operational burden and costs, through downsizing and merging processes. In this context, Italy has been interested by a progressive merging process of educational institutions. Indeed, while the highly centralized system originally allowed the functioning of diversified institutions on the basis of the offered educational level (for instance, kindergartens, primary schools, junior secondary schools), recently a new regulation has been adopted, oriented at promoting a higher degree of autonomy of educational institutions. However, in order to benefit from autonomy, schools have to comply with a series of requirements. In particular they must have a students' population between 500 and 900 units (that can be reduced to 300 in special cases, including islands and municipalities in mountain areas) that has to be demonstrated stable in this range for the last 5 years. If these requirements are not satisfied, schools must merge themselves with other institutions (belonging to any of the three categories) in order to form clusters that should

doi:10.1016/j.sbspro.2013.12.823

have a minimum students' population of 1000 units (reduced to 500 in the above-mentioned special cases) and include institutions from each educational level. This merging strategy allows schools rationalizing administrative and management offices, coherently to the cited policies of reduction of public expenditure. In practice, this process should be implemented by grouping schools in clusters, and letting each cluster being managed (in a centralized way) through the definition of a single cluster centre, providing shared administrative and managerial services. In this context, the availability of tools and models to assess and implement clustering (for grouping schools together) and locational choices (for locating clusters centers) could be useful for Local Authorities engaged in decision making activities that may provide cost savings and minimize, at the same time, the worsening of the service level for users.

In the operational research literature, several authors studied problems related to the organization of school systems; a typical problem concerns the so-called school districting, i.e. the partitioning of the demand coming from a given region in groups of students attending each school. In this problems school and class capacity constraints must be satisfied, various social objectives have to be achieved (for instance, racial balance) and some territorial aspects related to the contiguity of districts have to be considered to allow students from the same neighborhood to be assigned to the same school. The problem also occurs when reorganization actions have to be planned, such as the opening or the closing of a school, the modifying of the capacities of existing schools. In these cases, a perturbation of the previous demand allocation occurs, therefore the districts have to be redesigned considering a potential worsening of the accessibility of users to the service. In this context many models and methods have been defined and applied (see, for instance: Ploughman et al., 1968; Holloway et al., 1975; Brown, 1987; Lemberg and Church, 2000; Caro et al., 2004).

However, compared to classical school districting cases, the reorganizational problem faced by Italian institutions can be considered a more strategic problem, as it involves the organizational structure of the school system. Indeed, while the first class of problems concerns the assignment of students to schools (existing or to be located) in order to minimize a certain cost or distance function, in the second problem existing schools, with the related students' population, have to be grouped together under a shared management centre, in order to improve the efficiency in terms of operating costs. These problems can be addressed by using adaptation of very well established tools in the operational research literature, namely Facility Location models.

A facility location problem is aimed at finding the best position for a set of facilities within a given region in order to optimise a specific objective function (Daskin, 1995; Drezner and Hamacher, 2002; Eiselt and Marianov, 2011), in presence of a potential and/or actual demand that needs to be allocated to the facilities themselves on the basis of prescriptive rules (Hodgson, 1978; Beaumont, 1980) or of an utility function (see for instance: Bucklin, 1971; McLafferty, 1988; Lowe and Sen, 1996; Drezner and Drezner, 2001; Bruno et al., 2010). In particular Bruno and Improta (2008) and Bruno and Genovese (2012) used utility functions to show the mechanism of choice of students of their own preferred university site.

Historically, facility location models have been a viable decision support tool for institutions and firms that are planning to open new facilities or expanding their capacity in a given region. However, recently, in order to cope with practical situations emerging in the above-mentioned context of economic crisis, demand uncertainty and cuts to public expenditures, in the academic literature, several approaches have been proposed to solve downsizing, relocation and rationalization problems in a variety of scenarios (including both public and private sectors) through adaptations of traditional location analysis models. For instance, Wang et al. (2003) introduced a model addressing the situation in which, due to some occurring changes in the distribution of the demand, the relocation of the existing facilities in the location space is required in order to improve the service level provided to users. This approach simultaneously considered the opening of some new facilities and the closure of some existing ones. ReVelle et al. (2007) introduced the *Planned Shrinkage Model* that explores the reduction of facilities in a region. Starting from a common framework, they formulated two different models suitable, respectively, for a competitive and a non-competitive environment. Conversely, the p-median problem under uncertainty (Berman and Drezner, 2008) is aimed at locating, initially, p facilities, knowing that up to q additional facilities will be located in the future due to changes in demand.

Linking to this body of literature, in this work we present a location model aimed at describing and solving the school clustering problem as defined in the Italian case. The model is oriented to identify, within a given location space, the set of school facilities that, if merged together in a cluster (and, therefore, sharing management and administration service), could improve the efficiency of the system; moreover, the model should also identify the best location for the cluster centers. While the reorganization of the system represents an opportunity for the planner to reduce costs, it may generally produce a detriment of the service quality offered to the users. For this reason, through the imposition of appropriate constraints, the model must provide solutions that represent a good trade-off between the goal of the decision maker and the need of the users. The proposed model has been tested on a real-world case and the obtained results have been shown and commented.

The paper is organized as follows. In the next section a mathematical model for the school clustering problem is introduced and described, while Section 3 describes its application to a real case study. Finally, conclusions and directions for further researches are drawn.

2. A Mathematical Model for the School Clustering problem

As mentioned above, the reorganizational process of the school system (i.e. the grouping of schools in clusters) is aimed at reducing management costs, as each cluster will be managed in a centralized way through the definition of a single cluster centre. In this case, the current position of schools is assumed to be fixed and no demand re-allocation will occur. Therefore the reorganization will not have a direct effect on users' accessibility to the service. However, if we consider only efficiency aspects, the solution provided by the model could be considered not desirable (or equitable) from users perspective (Marsh and Shilling, 1994; Eiselt and Laporte, 1995). For example, the dimension of clusters is an important factor to be taken into account, as over-dimensioned clusters could have some side effects on the complexity of the managerial structure and, therefore, on the service level offered to users. For this reason, the planner should find a trade-off solution between the need to minimize costs and the need of keeping the discomfort caused to users below a given threshold. The problem is inherently multi-objective; however, it can be also modeled by means of a single-objective mathematical programming formulation, in which one of the objectives is included in the model as constraint. In order to avoid mentioned organizational inefficiency, aspects related to the location, composition and dimension of each cluster have to be considered. In particular:

- *location* concerns the position of the cluster's centre, to be chosen among schools assigned to the same cluster;
- *composition*, concerns the type of schools to be included in each cluster;
- *dimension* is related to the students' population of each cluster. In addition to the minimum threshold required by governmental regulations, a limit on the maximum dimension could be taken into account, in order to obtain more balanced solutions.

If we assume to define the number p of clusters to be created, the problem consists of identifying the best position to assign to the clusters' centres and in the allocation of schools to each cluster.

Denoting with:

I	the set of nodes corresponding to the positions of each school;
J	the set of potential locations for clusters' centres ($J \subseteq I$);
K	the set of school types or levels ($K = \{1,2,3\}$, with $k = 1$ identifying the pre-primary level, $k = 2$ the primary level and $k = 3$ the lower secondary level);
l_{ik}	a binary label equal to 1 if and only if node i hosts schools of type k;
a_{ik}	the number of students of school type k at the school in i;
d_{ij}	the distance between nodes i and j;
y_j	a binary variable equal to 1 if and only if node j is a cluster's centre;
x_{ij}	a binary variable equal to 1 if and only if node i is assigned to the cluster with centre in j;

We can formulate the following mathematical model (*sizing model*):

$$\min \quad z = \sum_{i\in I}\sum_{j\in J} d_{ij}x_{ij} \tag{1}$$

$$x_{ij} \le y_j \qquad \forall i \in I, \forall j \in J \tag{2}$$

$$\sum_{j\in J} x_{ij} = 1 \qquad \forall i \in I \tag{3}$$

$$\sum_{j\in J} y_j = p \tag{4}$$

$$\sum_{i\in I} l_{ik}x_{ij} \ge y_j \qquad \forall\, j \in J, \forall k \in K \tag{5}$$

$$\sum_{i\in I}\sum_{k\in K} a_{ik}x_{ij} \ge N_{min}y_j \qquad \forall j \in J \tag{6}$$

$$y_j \in \{0/1\}; x_{ij} \in \{0/1\} \qquad \forall i \in I, \forall j \in J \tag{7}$$

The objective function (1) represents the average distance between schools and their assigned cluster's centre, to be minimised. This objective represents one of the classical compactness measure used in the literature related to districting and clustering models (Ricca et al., 2011). Constraints (2), (3), (4) are classical p-median constraints, ensuring that: (2) node i is assigned to j only if node j is a cluster's centre; (3) each node i is assigned to only one cluster; (4) the number of clusters is equal to p. Conditions (5) impose that in each cluster j there is at least one school of each level k. Conditions (6) ensure that each cluster j has a students' population higher than N_{min}. Constraints (7) define the nature of decision variables.

It should be highlight that the presence of constraints (6) could lead to solutions characterized by a very skewed distribution of students' population among the produced clusters. In order to take into account balancing objectives, an additional set of constraints on the maximum population for each cluster has to be considered as follows:

$$\sum_{i\in I}\sum_{k\in K} a_{ik}x_{ij} \le N_{max} \qquad \forall j \in J \tag{8}$$

The introduction of constraints (8) would limit the maximum dimension of a cluster and produce more balanced solutions with higher values of objective function. In order to support the decision maker in the choice of trade-off solutions between the objective function and balancing aspects, it could be possible to perform a sensitivity analysis in terms of N_{max}, by varying this parameter between a lower bound $LB_{N_{max}}$ and an upper bound

$UB_{N_{max}}$. Indeed, decreasing N_{max}, it is possible to evaluate the trade-off between the objective function (expressing a rationalization need) and the balancing constraint.

As regards the upper bound $UB_{N_{max}}$, it is represented, for each value p, by the maximum dimension of clusters obtained solving the model (1-7); while an ideal lower bound can be assumed equal to the average students population for each group ($\bar{a} = \frac{1}{p}\sum_i \sum_k a_{ik}$). However, this value does not take into account the integrity constraints imposing that the population of each school has to be entirely assigned to the same cluster and the constraints on the composition of the groups. In order to calculate a more reliable lower bound value, a partitioning model was devised. Using the notation introduced above, the model can be formulated as follows:

$$\min \quad z = N_{max} \tag{9}$$

$$\sum_{j=1}^{p} x_{ij} = 1 \qquad \forall i \in I \tag{10}$$

$$\sum_{i \in I} l_{ik} x_{ij} \geq 1 \qquad \forall\, j \in \{1, \dots, p\}, \forall k \in K \tag{11}$$

$$\sum_{i \in I} \sum_{k \in K} a_{ik} x_{ij} \leq N_{max} \qquad \forall\, j \in \{1, \dots, p\} \tag{12}$$

$$x_{ij} \in \{0/1\} \qquad \forall i \in I, \forall\, j \in \{1, \dots, p\} \tag{13}$$

The objective (9) consists in the minimization of the maximum cluster size N_{max}. Constraints (10) indicate that each school i can be assigned to exactly one of the p clusters. Constraints (11) assure that, for each cluster j, there is at least a school i of level k, while conditions (12) state that the dimension of each cluster cannot exceed the value N_{max}. Constraints (13) concern the binary nature of the decision variables.
For each value p, the solution provided by this model represents the most balanced partition of the set of the schools in p groups; for $N_{max} \leq LB_{N_{max}}$, it is not possible to have feasible clusters.

The proposed model (and its variant including constraints (8)) will be tested on a real-world case study in the next section. This model can be solved by varying the parameter N_{max} between a lower bound $LB_{N_{max}}$ and an upper bound $UB_{N_{max}}$. In particular, $LB_{N_{max}}$ can be obtained solving the model (9)-(13); as regards the upper bound $\mathrm{UB}_{\mathrm{N}_{\mathrm{max}}}$, it is represented, for each value p, by the maximum dimension of clusters obtained solving the model (1-7).

3. The case study

The case study is focused on the aggregation of school institutions related to an urban district in the Municipality of Naples (about 12 Km^2) with more than 100.000 inhabitants. In this area there are 29 schools of different levels with a total number of 9077 students. The distribution of students is characterized by significant differences across the schools (ranging from a minimum of 40 students to a maximum of 798). The current arrangement is based on 11 clusters grouping all the schools. Figure 1 shows the position of each school and their aggregation in clusters. Table 1 indicates the composition of each cluster with the indication of students' population of each school. It has to be highlighted that the current organization does not satisfy governmental requirements both on the minimum students' population of clusters and on their composition, as most of the clusters do not include schools of each level (as prescribed).

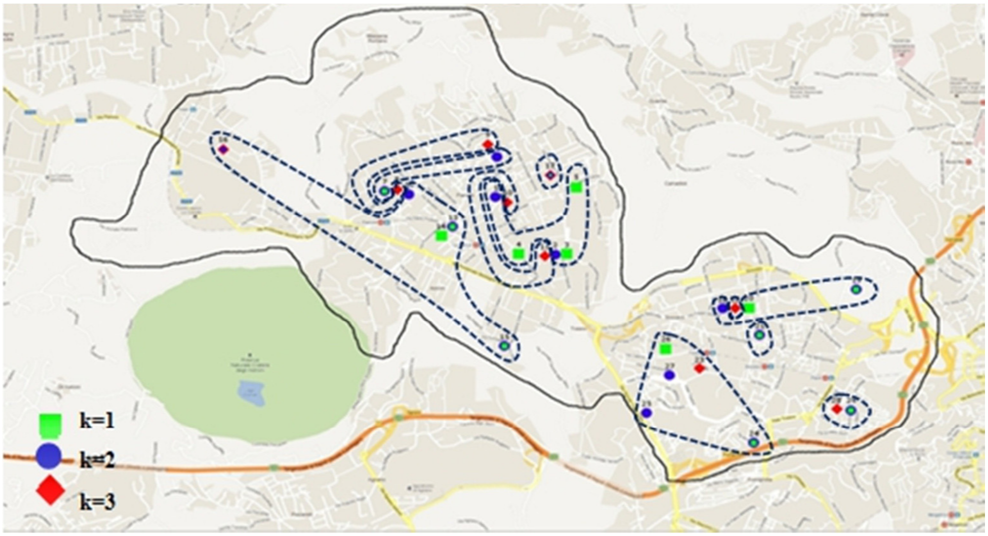


Figure 1. Location of schools and current organization in clusters

Table 1. Students' population data

Clusters	Schools *i*	Levels *k*			Total Population
		1	**2**	**3**	
I	1	-	433	-	**968**
	2	-	239	-	
	3	88			
	4	168			
	5	40	-	-	
II	6	-	601	-	**1202**
	7	277	324		
III	8	-	-	730	**880**
	9	-	-	150	
IV	10	-	-	770	**990**
	11	-	-	220	
V	12	-	474	-	**1025**
	13	100	66	-	
	14	155	-	-	
	15	20	75	-	
	16	-	96	39	
VI	17	76	80	100	**256**
VII	18	-	424	-	**749**
	19	55	138	-	
	20	132	-	-	
VIII	21	229	569	-	**798**
IX	22	-	-	700	**700**
X	23	-	-	203	**647**
	24	64	110	-	
	25	-	79	-	
	26	43	-	-	
	27	-	148	-	
XI	28	-	-	358	**862**
	29	160	344	-	

The proposed model has been applied to the case study and optimally solved using CPLEX 12.2 on an Intel Core i7 with 1.86 GigaHertz and 4.00 GigaBytes of RAM. Running times to obtain solutions are very limited (few seconds).

In the following, results are illustrated and discussed. In particular, the basic version of the model (including constraints 1-7 and fixing N_{min} =1000 and *I=J*) was considered first; then, the balanced one was evaluated, with the addition of constraints (8). Table 2 indicates the results obtained by varying the number p of clusters to be created from 2 to 7. For each solution, the number of students for each school (in decreasing order), the average and maximum distances between schools and the related cluster centers are reported.

As expected, in the case of the sizing model, the objective function decreases over *p*. The range of the size of each cluster, in general, also tends to decrease, even if this condition is not assured. For example, considering the passage from *p=5* to *p=6*, even if the average distance decreases of 17.18% (from 0.75 to 0.64 km), the solution for *p=6* is less balanced as the minimum dimension decreases from 1371 to 1055.

Table 2. Sizing model, N_{min} = 1000

Number of clusters *p*	**2**	**3**	**4**	**5**	**6**	**7**
Average distance	1.28	1.02	0.87	0.75	0.64	0.60
Maximum distance	3.05	3.05	3.05	2.50	2.50	2.50
Students for each cluster	5321	3877	3877	2411	2411	1466
	3756	3756	2247	2247	1539	1445
		1444	1509	1539	1445	1444
			1444	1509	1371	1371
				1371	1256	1256
					1055	1055
						1040

As above mentioned, the second version of the model has been implemented by varying the parameter N_{max} between a lower bound $LB_{N_{max}}$and an upper bound $UB_{N_{max}}$. In the Table 3 the bounds of the feasible range $\left[LB_{N_{max}}, UB_{N_{max}}\right]$, for each value of p, are reported.

Table 3. Feasible ranges for N_{max}

	p=2	*p=3*	*p=4*	*p=5*	*p=6*	*p=7*
$LB_{N_{\max}}$	4539	3026	2270	1816	1513	1297
$UB_{N_{\max}}$	5321	3877	3877	2411	2411	1466

The model including constraints (1-8) has been implemented, for each value of *p,* by varying $\mathrm{N_{max}}$ in the range $\left[LB_{N_{max}}, UB_{N_{max}}\right]$, with the following discrete step:

$$\Delta = \frac{UB_{N_{max}} - LB_{N_{max}}}{10}$$

such that:

$$N_{max} = LB_{N_{max}} + k_1 * \Delta, \text{ with } k_1 = 0, \dots, 10$$

The comparison of the solutions provided by the two models is summarized in Figure 2, where, for each p, the value of the average distance between schools and their respective cluster by varying the value of N_{max} is reported. In order to interpret the results, a single curve (for a given value of $p = \bar{p}$) can be considered. It is possible to notice that for $N_{max} = LB_{N_{max}}(\bar{p})$ we obtain a perfectly balanced solution with the maximum value of objective function. Increasing N_{max}, so gradually relaxing the balancing condition, the model provides better results in terms of objective function. In particular the minimum value is obtained when $N_{max} = UB_{N_{max}}(\bar{p})$. Of course for $N_{max} > UB_{N_{max}}(\bar{p})$, this does not vary anymore, as constraints (8) are not active. Coherently in Figure 2, for a given value of $p = \bar{p}$, the shape is monotonically decreasing until the value of the upper bound is reached. Of course this last conditions is obtained for lower values of N_{max} as p increases.

Furthermore, Figure 2 can have interesting managerial implications; in fact it can support the choice of the triplet (p, N_{max}, z). In particular, the graph can be interpreted in two ways: by fixing a value of objective function or, alternatively, by fixing a specific maximum value of clusters. The first interpretation consists in drawing a horizontal straight line ($z = f_0$) that allows the identification of:

- the values of p that ensure the achievement of that value of objective function;
- the corresponding values of maximum number of clusters.

In this case, it could be possible to identify the most preferable combination $(p, N_{max},)$.) for achieving the defined objective function value. The second interpretation consists in drawing a vertical straight line ($N = N_{max}$) that allows identifying:

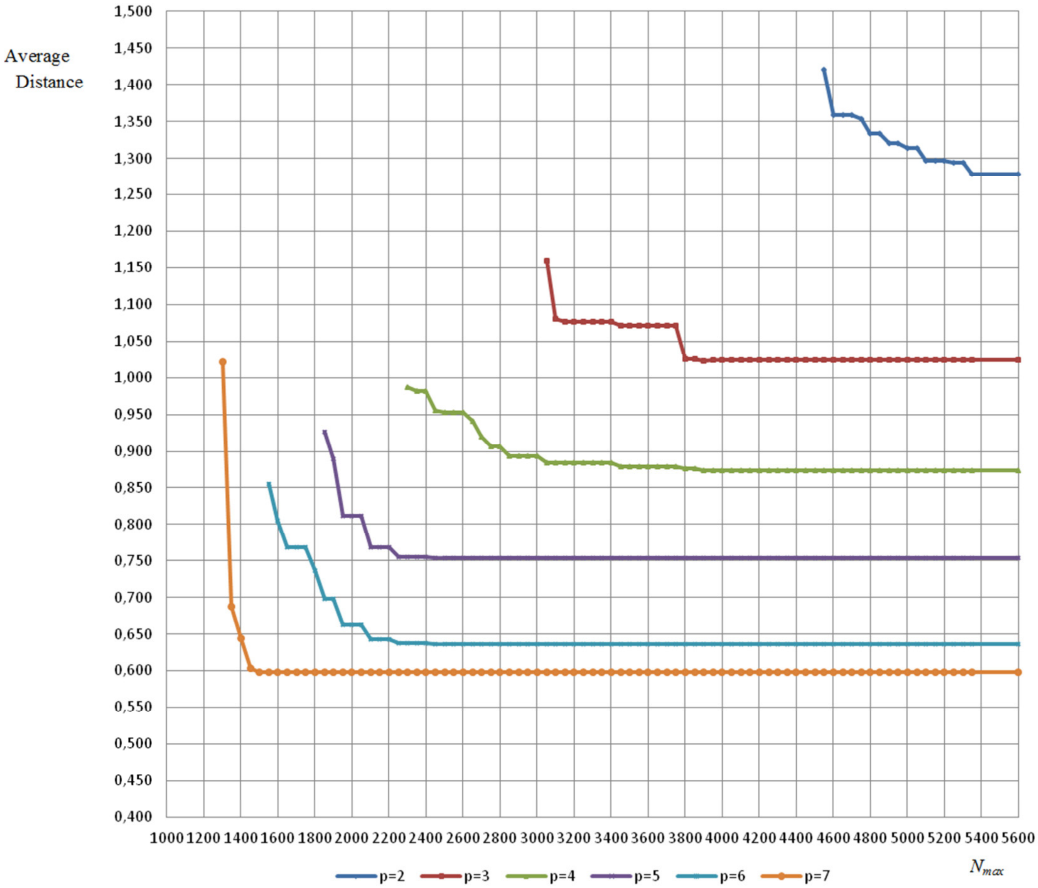
- the values of p that allow satisfying the constraint on N_{max};
- the corresponding objective function values.

In this case, it could be possible to compare the possible combinations (p, z) for forming clusters of the given maximum size. For example, considering $\mathrm{N_{max}} = 2400$, it can be understood that at least 4 clusters have to be formed. From Table 4, it can be seen that the objective function value improves in a very significant way increasing p from 4 to 5. It can be then derived that further improvements (obtainable by considering an increased number of clusters) follow a diminishing returns law; therefore, this would suggest that opening more than 5 school clusters might not produce improvements that can justify the increase in costs, as *p*=6 *and p*=7 return very small improvements in the objective function.

Table 4. Objective function variation for N_{max} = 2400.

	p=4	**p=5**	**p=6**	**p=7**
z	28.48	21.90	18.51	17.35

Figure 2. Average distance in function of N_{max}



4. Conclusions

In this work we analyzed a clustering problem concerning the reorganization of schools located in a given location space. In particular, the problem has been tackled considering the requirements indicated for the Italian case related to the reorganization of kindergartens, primary and junior secondary schools in institutions managed in a centralized way. In order to solve the problem, we proposed two versions of a mathematical model in dependence on the presence of balancing constraints. The models have been solved considering a real case study and the presentation of the results shows how they can be effectively used as decision support system. Further development of the research could include the application to different cases and the extension of the model to describe similar problems concerning the reorganization of different categories of public services (for instance, healthcare and administrative systems).

References

Beaumont J.R. (1980). Spatial interaction models and the location-allocation problem. *Journal of Regional Science*, 20:37-50.

Berman O., Drezner Z. (2008). The p-median problem under uncertainty. *European Journal of Operations Research*, 189:19–30.

Brown P.J.B. (1987). Trade-off and compromise in secondary school reorganization schemes: a goal programming approach. *Papers in Regional Science*, 61:115-130.

Bruno G., Genovese A. (2012). A spatial interaction model for the representation of the mobility of university students on the Italian territory. *Networks and Spatial Economics,*12(1):41-57.

Bruno G., Genovese, A., Sgalambro, A. (2010). An Agent-Based framework for modelling and solving location problems. *TOP*, 18(1):81-96.

Bruno G., Improta G. (2008). Using gravity models for the evaluation of new university site locations: A case study. *Computers and Operations Research*, 35 (2): 436-444.

Bucklin L.P. (1971). Retail gravity models and consumer choice: a theoretical and empirical critique. *Economic Geography*, 47:489-497.

Caro F., Shirabe T., Guignard M., Weintraub A. (2004). School redistricting: embedding GIS tools with integer programming. *Journal of the Operational Research Society,* 55:836-849.

Daskin M.S. (1995). Network and discrete location: models, algorithms, and applications. Wiley, New York.

Drezner T., Drezner Z. (2001). A note on applying the gravity rule to the airline hub problem, *Journal of Regional Science*, 41:67-73.

Drezner Z., Hamacher H.W. (eds) (2002). Facility location: application and theory. Springer-Verlag, Berlin.

Eiselt H.A., Laporte G. (1995). Objectives in locations problems, in Drezner Z. (ed), Facility location: a survey of applications and methods, pp. 151–180, Springer-Verlag, Berlin

Eiselt H.A., Marianov V. (eds) (2011) Foundations of Locational Analysis. Springer-Verlag - Berlin.

Hodgson M.J. (1978). Toward more realistic allocation in location allocation models: an interaction approach. *Environmental Planning*, 10:1273-1285.

Holloway C.A., Wehrung D.A., Zeitlin M.P., Nelson R.T. (1975). An interactive procedure for the school boundary problem with declining enrollment. *Operations Research*, 23:191-206.

Lemberg D.S., Church R.L. (2000). The school boundary stability problem over time. *Socio-Economic Planning Sciences,* 34:159-176.

Lowe J.M., Sen A. (1996). Gravity model applications in health planning: analysis of an urban hospital market. *Journal of Regional Science*, 36:437-461.

Marsh M.T., Schilling D.A. (1994) Equity measure in facility location Problem: A review and framework. *European Journal of Operational Research*, vol.74, pp. 1-17.

McLafferty S. (1988). Predicting the effect of hospitals closure on hospital utilization patterns. *Soc Sci Med*, 27:255-262.

Ploughman T., Darnton W., Heusser W. (1968). An assignment program to establish school attendance area boundaries and forecast construction needs. *Socio-Economic Planning Sciences,* 1: 243-258.

ReVelle C, Murray A.T, Serra D. (2007). Location models for ceding market share and shrinking services. *Omega, the International Journal of Management Science*, 35: 533-540.

Ricca, F., Scozzari, A., Simeone, B. (2011). Political districting: from classical models to recent approaches. *4OR*, 9:.223-254.

Sancton, A. (2000). Merger Mania: The Assault on Local Government. McGill-Queen's Press.

Wang Q., Batta R., Badhury J., Rump C. (2003). Budget Constrained location problem with opening and closing of facilities, *Computers and Operations Research*, 30:2047-2069.