

A Geographical Information System Framework for Evaluating the Optimum Location of Point-Like Facilities

Anastasios Karaganis and Angelos Mimis
Department of Economic and Regional Development,
Panteion University of Social and Political Sciences, Sygrou Avenue 136, 17671 Athens, Greece

Abstract: This study discusses a methodology for solving the problem of finding the optimum location of a number of facilities entering an existing network. The locational optimization problem is approached with a continuous model and the facilities are considered to be point-like. The methodology proposed solves these problems by using Voronoi diagrams to define the catchment area and minimize the average distance travelled by the users. The approach is implemented by loose coupling between the Directed Tabu Search algorithm for continuous non-linear optimization with constraints and the Geographical Information system where the cost function is evaluated. Two different case studies are illustrated. In the first, where point-like facilities are used by independent users, two new post offices are inserted in a sub region of Athens where a number of post offices are already in use and gave an improvement of 40% in the cost function. In the second case, facilities operate in a hierarchical network comprising by bank branches and automated teller machines of a private Greek bank. A new bank branch location was examined and compared with the insertion of an Automated Teller Machine and provided an improvement of 25.6 and 12.8%, respectively.

Key words: Geographic information systems, spatial analysis, non-linear programming, locational planning, private Greek bank, network

INTRODUCTION

Locational optimization problems have been studied for >100 years. They dated back to Launhardt in 1882 and to Weber in 1909 (Puu, 1997). In their more general form, they require as an input the study area, the metric used to evaluate distances and the density of users in the area of interest (ReVelle and Eiselt, 2005). As a result, they are used in the calculation of the optimum location of a number of facilities providing specific services based on a cost function (Plastria, 2001). The variety of applications is wide, starting from location for fire stations, school units, bus stops up to location of distribution centers, retail market and sport fields.

The methodology that has been presented in literature for solving locational optimization problems can be classified into four major classes based on the way the region of the problem is treated and the cost function (ReVelle *et al.*, 2008).

The first class captures the analytic models which are based on a number of simplifications of the problem such as the hypothesis that demand rises homogeneously over the region. These assumption lead into obtaining closed

form solutions which are function of the parameter of the problem such as the number of facilities to be inserted. Despite that these models give a clear picture of the relation between the location of the facilities and the parameters of the problem, the assumptions used make it difficult to adopt it in real life applications (Leamer, 1968; Geoffrion, 1979).

In the second class are included network models where the problem of facilities location is treated in a network consisting from links and nodes. In most applications of these models, the demand appears only in nodes and an exception to that is the demand of emergency services in the road network where demand appears in links as well as in nodes (ReVelle *et al.*, 2008).

Another class is the discrete location models where demand and supply of a service appear in discrete points in the study region. Their mathematical formulation leads into N-P hard problems of integer programming. These models are the most common in literature and they come in different formulations such as the p-median model, the location set covering model and the center model. In the first case the aim is to minimize the demand-weight average distance. In the next case the aim is to minimize

the number of facilities that are needed to provide service to all users. In the last model the aim is to minimize the maximum facility-customer distance (ReVelle *et al.*, 2008; ReVelle and Eiselt, 2005).

The last class is the continuous models which assume that facilities can be located anywhere in the study region and the demand arises usually in a discrete number of points. These problems are treated by methods of non-linear programming and algorithms of finding global optimum. Despite the critics relating to its inability to model real life, it has been applied to a number of problems as the location of mail boxes in transportation and in logistics (Novaes *et al.*, 2009).

The continuous models with the use of Voronoi diagrams are solving three different kinds of problems which are either point-like, line or space-time problems (Okabe, 2000). In the first case where point-like facilities are treated, the goal is to find the optimum location in order to minimize the average cost. The point like facilities can be facilities that are used by independent users or by groups of users, hierarchical services, observation points, service points of mobile facilities.

It is evident that surveys using a connection between the location models, the Geographical Information System (GIS) and efficient algorithms able to solve these problems are rare and problem specific (Bozkaya *et al.*, 2010; Straitiff and Cromley, 2010). Example of such research is a GIS framework, proposed by Bozkaya *et al.* (2010) of an integrated location-routing model and a hybrid heuristic solution methodology for a competitive multi-facility location problem in the Istanbul area.

The current study tries to fill this gap by loosely coupling the GIS with a continuous location model based on Voronoi diagrams and the global optimization algorithm named Directed Tabu search.

Description of the problem: In the continuous location model adopted the facilities can be located anywhere in the service area and demand arises in discrete points. Two location problems will be discussed (Okabe, 2000; Okabe and Suzuki, 1997). In the first one the problem of finding the optimum location of point facilities which are used by independent users will be formulated. Example of such service is the location of mail boxes. In the second case the optimum location of hierarchical facilities will be discussed. An application of such a model is the main library and its branches (Okabe *et al.*, 1997).

Optimum location of point facilities used by independent users: In order to examine quantitatively the proposed location, a measure of comparison is needed or in other words a cost function. In the problem discussed two

sources of cost arise. The first cost is related to the location and the second is born by the users of the service. In the current research only the second part of the cost is modeled by assuming that the aim is to minimize the user's cost. This is the case especially in public services.

In that respect the cost function can be defined in terms of the distance that the users are travelling in order to move into their nearest service. So every point-like facility has a catchment area which can be defined in terms of Voronoi polygons. As a distance function the Euclidean distance is used which is close to reality as the comparisons have shown (Apparicio *et al.*, 2008; Love *et al.*, 1988). Thus, the total cost involved is the mean distance that the user will travel to its closest facility.

In order to formulate mathematically the problem, assume that $S \subset R^2$ is the region of interest and:

$$x_1, x_2, \dots, x_n \in S$$

the location of n point facilities. If $\phi(x)$ is the density of users in the region S and:

$$V_1, V_2, \dots, V_n$$

the Voronoi polygons that is related the point-like facilities. So the cost function is given by:

$$F(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \int_{V_i} f(\|x - x_i\|) \phi(x) dx \quad (1)$$

where, $\|x - x_i\|$ is the Euclidean distance. If:

$$f(\|x - x_i\|) = c \|x - x_i\|^2$$

the problem of optimum location can be written as:

$$\min \sum_{i=1}^n \int_{V_i} \|x - x_i\|^2 \phi(x) dx \quad (2)$$

with the constraint that the location of point facilities:

$$x_1, x_2, \dots, x_n$$

lies within the region S. The similar network problem is the multi-median problem.

Optimum location of hierarchical facilities: Consider n locations of a service in region S that are ranked from 1 to m. Assume that n_i of them have rank i:

$$\left(\sum_{i=1}^m n_i = n \right)$$

and x_1, x_2, \dots, x_n are the location of n facilities. Further suppose that the first n_1 have rank 1, the next n_2 have rank 2, etc. Further assume that n facilities supply k different services S_1, S_2, \dots, S_k and facilities of rank i supply S_i, S_{i+1}, \dots, S_m services or in other words the service of its rank and all the services of its lower rank $i+1, i+2, \dots, m$. This is known as successively inclusive hierarchy.

Before going into the mathematical formulation of the cost function, two assumptions will be made. First every user goes to its nearest facility that provides the required service. So the catchment areas are given by dividing the region, for every rank, into Voronoi polygons. Secondly every user consumes a_j time units of service s_j . Finally it is assumed that:

$$\sum_{i=1}^k a_i = 1$$

So the cost function is:

$$F(x_1, x_2, \dots, x_n) = \sum_{i=1}^k a_i \sum_{j=n_1+\dots+n_{i-1}+1}^n \int_{V_j^i} (\|x_j - x\|) \phi(x) dx \tag{3}$$

As it can be noticed the two formulations in Eq. 3 and 1 are similar and can be solved with the same software by merely altering the cost function.

MATERIALS AND METHODS

In order to solve the problems of optimum location of facilities that has been described in the study, software that incorporates two evaluating mechanisms is needed. The first is a Geographical Information System (GIS) where with the usage of the relating spatial information, the cost function could be evaluated by using specific developed code. The second is an optimization algorithm, enabling us to solve non-linear problems with constraints. In the absence of an integrating solution in the GIS, a loose coupling approach was used to integrate these two mechanisms. The first part is based on Mapbasic scripting language of MapInfo (2008) and the optimization algorithm was run on Matlab.

The optimization problems are not new in the GIS environment (Van Dijk *et al.*, 2002). The problems of map labeling, generalization and line simplification have

already been treated by global optimization techniques such as Genetic algorithms and modern heuristic search techniques. These methods are in demand especially because they do not have to incorporate problem assumptions and find near-optimal global solutions fast. In this case the method of optimization adopted is a flavor of Tabu search (Glover, 1989, 1990) for continuous fields (Directed Tabu Search) that has been developed by Hedar and Fukushima (2006). The method Directed Tabu Search (DTS) that is used is based on the metaheuristics Tabu search method that has been enhanced so that it can treat problems of continuous functions. The DTS uses direct-search methods which provide the direction of search for the Tabu method. These techniques are based on the method of Nelder-Mead (Kelly, 1999) which has been enriched by the adaptive pattern search. In this part of the framework and before going into evaluating the cost function, the problem parameters need to be defined. These are the number of facilities that need to be inserted in the study region. It should be reminded also that for each facility, the problem variables are two, one for each coordinate. Another important point is the constraints having the aim to keep the optimum locations within the study area. These are implemented at first as inequality constraints in the DTS, equal to min-max box of the study region and in a second phase as penalty functions.

The cost function is evaluated in the commercial GIS package MapInfo, for the problem parameters (location and number of facilities) that are required by the optimization procedure. The script for the calculation is developed in the language MapBasic (MapInfo, 2008). The cost functions of the two problems modeled are described by Eq. 1 and 3. Specifically in the first case the cost function takes the discretized form:

$$F(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^{bl(i)} \frac{Area_{V_i}}{Area_{total}} \cdot \frac{Population^j}{Population^{total}} \cdot \|x - x_i\| \tag{4}$$

Where:

- x_1, x_2, \dots, x_n = The location of n point-like location facilities
- $Area_{V_i}$ = Area of V_i polygon
- $Area_{total}$ = Total area of study region
- $\|x - x_i\|$ = Distance of j block from the x_i facility
- $Population^j$ = Population in j block
- $Population^{total}$ = Total population in the study area

As it can be seen from Eq. 4 the cost function is evaluated by summing up, for each Voronoi and for each

block within the distance that the customers have to travel to the facility, weighted by the percentage of the area of this block to the total area and by the percentage of the population of this block to the total population. Similarly, one can derive the second discretized form of Eq. 3.

RESULTS AND DISCUSSION

Based on the methodology described above and on the tool developed for that purpose, two applications will be presented. In the first one the problem of finding the optimum location of new post offices on a network that is already in use. In the second application the problem of locating either a new bank branch or a new Automated Teller Machine (ATM) on an existing hierarchical network of a private Greek bank is considered. The study area is a sub region of Athens consisting by 5 municipalities has a total area of 21.4 km² and total population of 375.000. Also the population is known and used for each block (Fig. 1).

Optimum location of post offices: On a previous survey using the same location model and assumptions, the optimum location of mail boxes has been investigated in a rectangular grid (Tokyo area) where all the facilities were inserted at the same time (Okabe, 2000). On the contrary in the case the study area is a non-convex sub region of Athens as shown in Fig. 1 were facilities are already in use.

In this case the problem is treated assuming that the services are used by independent users and for this purpose, nine post offices have been geocoded in the study area. The cost function is given by Eq. 2 and the population density $\phi(x)$ is evaluated by using the layer of population.

The current implementation can give answers to questions such as where is the optimum position of m facilities in network of services already in use.

In Fig. 2, it is illustrated the solution of inserting two new facilities (post offices) in the study area where there are already 9 post offices in use. The cost function with the inclusion of a new post office, exhibits an improvement of 29% in comparison with the current network whereas the insertion of two extra facilities exhibits an improvement of 40%. It should be noted that in the solution of the problem with the algorithm DTS 807 function evaluations took place where each one of those needed 4 sec in an AMD Athlon x2250.

Optimum location of bank branches and ATMs:

Although, there various studies investigating the location of bank branches in a competitive environment (Miliotis *et al.*, 2002) and even merging or closure of bank branches (MacDonald, 2001; Morrison and O'Brien, 2001) or on the other hand location of ATMs (Aldajani and Alfares, 2009), there is no approach considering their interaction in a single model. In this case the problem of finding optimum location of bank branches and ATMs is modeled as a problem in a hierarchical network. For this

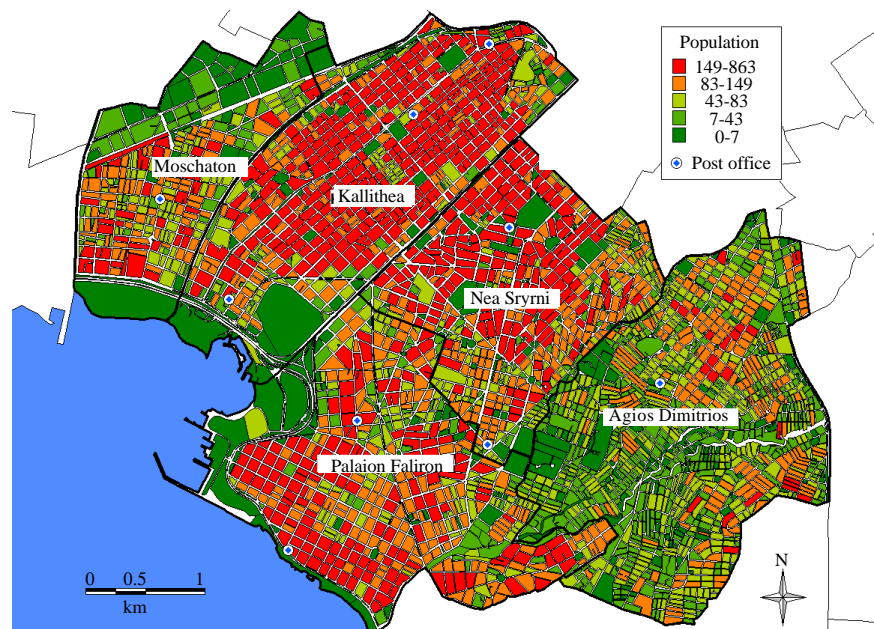


Fig. 1: The study area colored by population and the position of the post offices

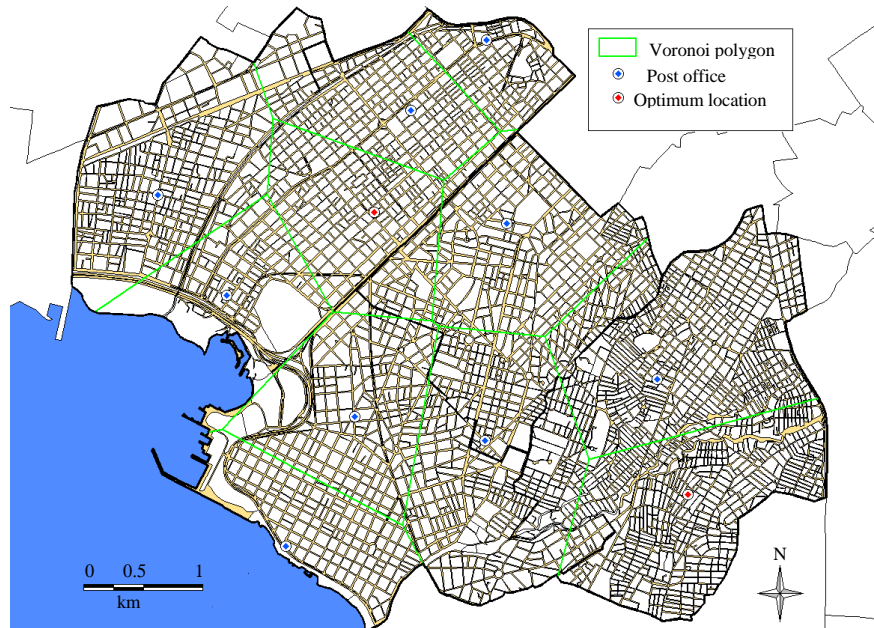


Fig. 2: The optimum position of two new post offices

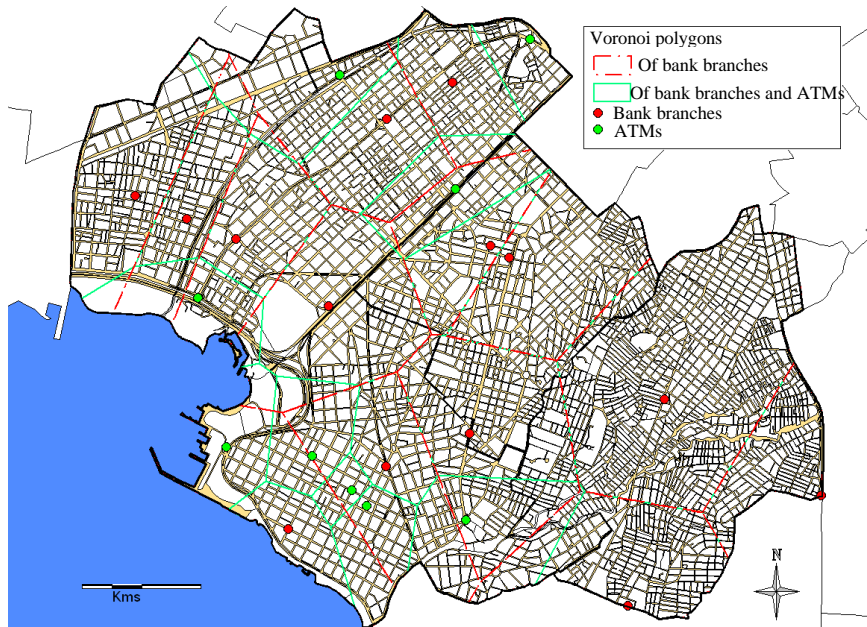


Fig. 3: The bank branches and the ATMs present in the study area

purpose a Greek bank was chosen and 14 branches and 4 ATMs were geocoded in the study area (Fig. 3). The bank branches were considered as being rank 1 and the ATMs as rank 2 based on the approach that all the services of ATMs and not limited to those are provided by branches as well. In this application as a cost function was used the Eq. 3 and the time units of usage are considered equal for each service i.e., $a_1 = a_2 = 0.5$.

This implementation provides us answers to questions such as where is the optimum location of n new bank branches and m ATMs in a network of services under use.

In Fig. 4, it is shown the solution of two problems. The first one refers to finding the optimum location of a new bank branch whereas the second to finding the location of an ATM in a network already in use. The

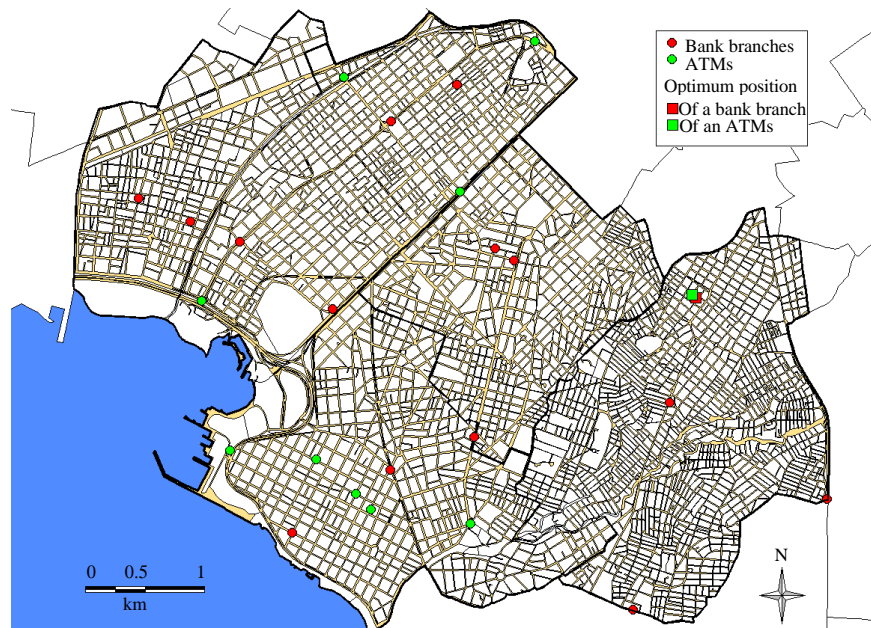


Fig. 4: The optimum position of a new bank branch and of a new ATM

insertion of the bank branch gave an improvement of 25.6% of the cost function and the insertion of an ATM gave a 12.8% improvement. In Fig. 4, it could be noticed that these two points are close. For the solution of these two problems 263 and 269 function evaluations were made, respectively and each function evaluation took about 5.4 sec.

CONCLUSION

In this study an approach of evaluating the optimum position of a new facility in a network in use is described. In this methodology the space is considered to be continuous and the services modeled are point-like facilities. Further the mathematical formulation of two problems was given. In the first one the facilities were used by independent users whereas in the second a hierarchical network was modeled.

Next, two applications were given, one for each formulation. In the first, two new post offices were inserted on a network already in use. The methodology gave an improvement of 40% in the cost function. In the second case, it was examined the insertion of a bank branch and of an ATM, alternatively, on a hierarchical network in use. The solution provided an improvement of 25.6 and 12.8%, respectively.

In this approach the Geographical Information Systems (GIS) was coupled with a location model based on Voronoi diagrams and an efficient global optimization algorithm. This resulted in displaying the consequences

of the user input on the map and most importantly, the changes that could occur from these modifications. This could be a useful in the hands of practitioners and provide new options in the decision making process (De Smith *et al.*, 2009).

It should be noted that real world problem are far more complex than the models presented since there are many other factors to be taken into account such as availability of suitable sites, cost of sites, convenience and regulatory issues. The limitations of the models adopted are based mainly on their characteristics, firstly because they are continuous and secondly because of the Euclidean distance that is used. The optimum sites found could be located in an unsuitable area e.g., occupied by another land use and be very costly to demolish and relocate. Also the use of straight-line distance does not take into consideration barriers present such as rivers, sea, mountains that may exist in the search area (Neema *et al.*, 2008). Despite that the location models can provide a helpful tool in the decision making process and can provide comparative alternative solutions.

REFERENCES

- Aldajani, M.A. and H.K. Alfares, 2009. Location of banking automatic teller machines based on convolution. *Comput. Ind. Eng.*, 57: 1194-1201.
- Apparicio, P., M. Abdelmajid, M. Riva and R. Shearmur, 2008. Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *Int. J. Health Geographics*, 7: 7-7.

- Bozkaya, B., S. Yanik and S. Balcişoy, 2010. A GIS-based optimization framework for competitive for multi-facility location-routing problem. *Networks Spatial Econ.*, 10: 297-320.
- De Smith, M.J., M.F. Goodchild and P.A. Longley, 2009. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. 3rd Edn., SPLINT, University of Leicester, UK..
- Geoffrion, A.M., 1979. Making better use of optimization capability in distribution system planning. *IIE Trans.*, 11: 96-108.
- Glover, F., 1989. Tabu search-Part I. *ORSA J. Comput.*, 1: 190-206.
- Glover, F., 1990. Tabu search-Part II. *ORSA J. Comput.*, 2: 4-32.
- Hedar, A.R. and M. Fukushima, 2006. Tabu search directed by direct search methods for nonlinear global optimisation. *Eur. J. Operat. Res.*, 170: 329-349.
- Kelly, C.T., 1999. *Iterative Methods for Optimization*. SIAM, Philadelphia, USA..
- Leamer, E.E., 1968. Locational equilibria. *J. Regional Sci.*, 8: 229-242.
- Love, R.F., J.G. Morris and G.O. Wesolowsky, 1988. *Facilities Location: Models and Methods*. North-Holland Publishing Company, New York.
- MacDonald, E.H., 2001. GIS in banking: Evaluation of Canadian bank mergers. *Can. J. Regional Sci.*, 24: 419-442.
- MapInfo, 2008. *MapBasic. Version 9.5. Reference Guide*. Pitney Bowes Software Inc., UK..
- Miliotis, P., M. Dimopoulou and I. Giannikos, 2002. A hierarchical location model for locating bank branches in a competitive environment. *Operat. Res.*, 9: 549-565.
- Morrison, P.S. and R. O'Brien, 2001. Bank branch closures in New Zealand: The application of a spatial interaction model. *Applied Geogr.*, 21: 301-330.
- Neema, M.N., K.M. Maniruzzaman and A. Ohgai, 2008. New genetic algorithms based approaches to continuous p-median problem. *Networks Spatial Econ.*, 10.1007/s11067-008-9084-5
- Novaes, A.G.N., J.E.S. de Cursi, A.C.L. da Silva and J.C. Souza, 2009. Solving continuous location-districting problems with Voronoi diagrams. *Comput. Operat. Res.*, 36: 40-59.
- Okabe, A. and A. Suzuki, 1997. Locational optimization problems solved through Voronoi diagrams. *Eur. J. Operat. Res.*, 98: 445-456.
- Okabe, A., 2000. *Spatial Tesselations: Concepts and Applications of Voronoi Diagrams*. 2nd Edn., John Wiley and Sons, New York, ISBN-13: 9780471986355.
- Okabe, A., K.I. Okunuki and T. Suzuki, 1997. A computational method for optimizing the hierarchy and spatial configuration of successively inclusive facilities on a continuous plane. *Location Sci.*, 5: 255-268.
- Plastria, F., 2001. Static competitive facility location: An overview of optimization approaches. *Eur. J. Operat. Res.*, 129: 461-470.
- Puu, T., 1997. *Mathematical Location and Land Use Theory*. Springer-Verlag, Berlin.
- ReVelle, C.S. and H.A. Eiselt, 2005. Location analysis: A synthesis and survey. *Eur. J. Operat. Res.*, 165: 1-19.
- ReVelle, C.S., H.A. Eiselt and M.S. Daskin, 2008. A bibliography for some fundamental problem categories in discrete location science. *Eur. J. Operat. Res.*, 184: 817-848.
- Stratiff, S.L. and R.G. Cromley, 2010. Using GIS and K = 3 central place lattices for efficient solutions to the location set-covering problem in a bounded plane. *Trans. GIS*, 14: 331-349.
- Van Dijk, S., D. Thierens and M. de Berg, 2002. Using genetic algorithms for solving hard problems in GIS. *Geoinformatica*, 6: 381-413.