Raport Badawczy

Research Report

RB/48/2015

The bi-partial versions of the k-means algorithm and their potential applications to the generic facility location problem

J.W. Owsiński

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
Dr hab. inż. Lech Kruś, prof. PAN

Warszawa 2015

The bi-partial versions of the k-means algorithm and their potential applications to the generic facility location problem

Jan W. Owsinski
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warszawa, Poland
owsinski@ibspan.waw.pl

Abstract

The present short report constitutes a continuation to the initial study of the bi-partial version of the well known original p-median or p-center facility location problem. The bi-partial approach, developed by the author, primarily to deal with the clustering problems, is shown here to work for a problem that does not, in principle, possess some of the essential properties, inherent to the bi-partial formulations. It is shown that the classical general objective function of the problem considered, namely that of facility location, can be correctly interpreted in terms of the bi-partial approach, that it in fact possesses the essential properties that are at the core of the bi-partial approach, and, finally, that the general algorithmic precepts of the bi-partial approach can also be effectively applied to this problem. In this particular case, the precepts of the classical k-means procedure are referred to. The effectiveness of the proposal is shown for certain variants of the problem, with the bi-partial counterparts appropriately defined. It is proposed that the use of bi-partial approach for the class of problems, represented by the general objective function of facility location, and its most basic variants, can be beneficial from the points of view of both flexibility and interpretation.

Keywords: facility location, p-median, p-center, clustering, bi-partial approach

1. The facility location problem

The facility location problem is among the classical, historically well-rooted problems of Operational Research, having its origins, in the sense, in which it is being formulated and solved nowadays, in the 19th century economic analyses. There are several, different, streams of thought, concerning the formulation, meaning, as well as the methods and sense of solving the problem. Those, who stick to its "geometrical", rather than economic, nature, refer much farther back than "just" to Alfred Weber, recalling Fermat and Torricelli from the 17th century and Simpson from the 18th century. This "geometrical" direction, though, is not of the primary interest here, even if it offers important insights, both in terms of numerical procedures, and the associated interpretations.

We wish to address the general facility location problem as it appears in most of the present-day studies, even though, of course, there are (numerous, indeed) aspects, appearing in some formulations (constraints, requirements) that shall not enter explicitly the form here considered. It is, namely, not our intention to demonstrate that the approach proposed can and should be used for all kinds of the so diverse facility location problems. That is why we address what we call the "generic facility location problem", i.e.

$$\min \Sigma_q(\Sigma_{i\in Aq}\, d(x_i,x^q) + c(A_q,x^q)) \quad (1)$$

where minimisation is performed with respect to $\{(A_q,x^q)\}_q$,

and where q is the index of the facility ($q = 1,\ldots,p$) and of its catchment (A_q), i is the index of the customer / demand locations ($i\in I = \{1,\ldots,n\}$), x_i being these locations and their characteristics, x^q, on the other hand – being the locations and characteristics of the facilities; the catchments A_q are interpreted as subsets of I, $A_q\subseteq I$. The functions $d(.,.)$ and $c(.,.)$ have the interpretation of costs, in the former case – costs closely associated with distance between x_i and x^q, while in the latter case – with, potentially, quite a variety of factors. These factors may include, in particular: (a) the setup cost (which itself may be constant, c, or, for instance, may depend upon the manner, in which I is split into A_q), (b) the capacity increase cost, depending directly upon cardinality of A_q, denoted n_q, (c) the location cost, $c(x^q)$, a simile of $d(.,.)$ for the upper-level supply problem (that is – depending upon the cost of transport from some external location).

The essential interpretation of (1) is that we wish to locate the "facilities" (x^q, $q = 1,\ldots,n$) and to establish their catchments A_q, i.e. we wish to determine $\{(A_q,x^q)\}_q$, given the set of customer / demand points x_i, that is – primarily their locations, in such a way as to minimise the overall cost (this cost may be understood in terms of [discounted] total cost of operation over the lifetime of the given undertaking, or, say annual cost, including amortisation of the capital expenses).

So, we wish to locate x^q so as to have the sums of costs-distances $d(x_i,x^q)$ for all x_i belonging to the catchments A_q as little as possible, while also minimising the cost of establishing and maintaining the facilities, $c(A_q,x^q)$. Were it not for the latter component, the "optimum" would be to have as many facilities as there are demand points, so that the absolute minimum of (1) is reached (possibly equal zero, see Section 2). At the other end of the stick just one, single facility would be located, (almost certainly) minimising the second cost component. Thus, we look for an equilibrium between the two components, this equilibrium corresponding to the minimum of (1).

Let us add at this point that formulation (1) implicitly – via the above mentioned notion of "equilibrium" – includes the determination of the optimum number of facilities / catchments, p. It must be emphasised, though, that in the majority of actual studies and models this number is simply given, such a far-reaching assumption being justified, on the one hand, by the algorithmic or numerical problems, and on the other – by the fact that in many cases the choice of the number and locations of the facilities is highly limited.

2. Some important details and their consideration

Demand variation across space

The demand points x_i may, of course, feature different demand volumes, and this fact ought to find a reflection in the problem formulation. In an explicit manner this could be:

$$\min \Sigma_q(\Sigma_{i\in Aq}\, w_i d(x_i,x^q) + c(A_q,x^q)) \quad (2)$$

where w_i correspond to demand volumes, so that, naturally, for higher w_i "shorter" distances $d(x_i,x^q)$ ought to be secured in the solution.

Obviously, given A_q and the locations of x_i in it, the best location for x^q is some sort of a "median", or a definite location that is possibly close to this "median", or to an otherwise defined centre, of this set of locations, provided distance $d(.,.)$ is appropriately defined, see further on. If so, the introduction of w_i would imply the search for a weighted "median" or centre, with w_i being respective weights.

It is, of course, possible to avoid the introduction of explicit weights by assuming some unit of demand, say w, and then placing at location x_i the number of demand points, corresponding to w_i/w. (Here, we leave out the question of what happens if w_i/w is not an integer or is not very close to an integer.)

In reality, the distinction of x_i and the association of demand w_i at these points is just the matter convenience, and not so much of modelling accuracy: it is most often the case that x_i actually correspond to some potential facility locations, to which demand, being generated more or less continuously over space, is simply assigned. Hence, we deal here with a regression issue, namely: how to determine the points x_i in an optimum manner, given density of demand over space, $w(x)$, and given "disturbances" to the continuity of space, in the form of, first of all, transport infrastructure.

Locations of facilities vs. locations of demand points

It is most natural and simplest to assume that we just deal with a set of locations (e.g. settlement units), corresponding to x_i, and that x^q are selected among these. The obvious examples are: the set I is the set of localities, where the facilities could be located, but where also demand is generated. If a facility x^q is located at x_i, then $d(x_i,x^q)$ equals zero (unless some special assumptions are made on the shape of $d(.,.)$), lowering the overall cost (especially when a high weight (demand) w_i is assigned to this location).

It is, however, quite imaginable that (a) not all of the locations of x_i are feasible from the point of view of locating x^q, and so we deal with a subset $I^Q \subseteq I$ of the potential locations of x^q; or (b) the set of potential locations of x^q is entirely different from that of x_i, although, of course, it has to belong to the same space from the point of view of distance calculation.

Actually, none of these three instances (exhausting, in fact, all of the possible ones), introduces a "qualitative" difference into the formulations (1) or (2) (limitations to the set of potential facility locations). The sole difference may lie, therefore, in the technical details of the respective algorithms.

Overlapping and/or fuzzy catchments

The function $c(A_q,x^q)$ is, in principle, meant to express the features of the facility q that directly depend upon the properties of the catchment A_q and the actual location in space (local demand and distances to cover). From the managerial point of view, though, a stiff and strict assignment of i to q may be an unnecessary constraint. Thus, demand from the locations i that belong to a neighbouring q might be satisfied from the second-closest, or even the third-closest facility q', when a need arises. The essential point is – to what extent is this accounted for in a potential solution?

If the choice of the second- or third-closest facility is just a matter of everyday managerial choice, then it is not necessary to account for this in the formulation of the problem and in the solution. Yet, a requirement may be formulated, bearing an influence on the form of the problem and the solution, that explicit rules be found of assigning demand locations x_i to facilities q, and thus the degree of complication may get much higher. Still, in the perspective that we aim at here, if the catchments A_q are defined as fuzzy sets of x_i, i.e. a location x_i may belong, in the solution, to the particular catchments A_q in the degrees $\mu_{iq} \in [0,1]$, with, possibly (and, indeed, reasonably), imposed condition that $\sum_q \mu_{iq} = 1$ for all $i \in I$ (this meaning that we deal with what is called the "Ruspini partition"), then, once found, the thus defined catchments can embody in quite a natural manner the requirements of the kind here mentioned.

Multi-location facilities

Another, similar requirement might concern the presence of several facility locations per catchment, like, e.g. "primary" and "secondary" centres, which serve slightly different purposes, or customers, for instance – depending upon the volume of goods or services (say, post-sales service network, with small outlets for either taking in the equipment for repair and giving it back to the customers, and higher rank facilities, where actual repairs are being done, parts are stored, etc.). In this manner, a "hierarchical" version of the problem might arise (see also the analogous comment on the regression issue, formulated in the preceding point) that would require determination of two levels of A_q, the lower level being denoted B_r, so that conditions

$$\cup_{q=1,\ldots,p} A_q = I = \{1,\ldots,n\}$$

and

$$\cup_{r=1,\ldots,\mathrm{card}A_q} B_{rq} = A_q, \ \forall\ q = 1,\ldots,p,$$

are preserved (the double index rq referring to the lower level catchments r belonging to the upper level catchments q). This requirement, indeed, changes the perspective on the solution, since, at least, the dimensionality of the problem increases dramatically, even if some of the essential features are preserved. Yet, despite this change of perspective, we shall return to this kind of problem in this report.

Extended versions: the universe of the actually solved problems

The aspects, which have been roughly outlined in the preceding points, constitute usually just a minor portion of the problems currently being formulated ("modelled") and solved in the broad domain, referred to as "location analysis". These problems include a variety of aspects, which, at least potentially, characterise the real-life situations (in addition to those already mentioned and expressed), such as, in particular:

-- limited facility capacities, usually associated with the value of $\sum_{i \in A_q} w_i$, when the weights w_i correspond to the transported / stored / produced quantity;

-- limited capacities of the edges (i,j), again meaning that the sum of w_i, coinciding in the solution with a given edge, is limited[1];

-- selection of transport means, that is – introduction of an index m, corresponding to the choice of transport means, with which various costs and capacities are associated;

-- accounting for the time dimension, for instance – in the form of succession of facility construction over time; or

[1] this aspect, like the following ones, cannot be accommodated directly in the formulations forwarded here – they require more elaborate forms of the problem, especially involving additional indices.

-- accounting for a more realistic cost calculation, including, for instance, cost of transporting a unit of goods, taking, in particular, into account the relation to the capacity of the transport means (including the negative impact of excess slacks).

All these lead to quite complex formulations, which are usually squeezed into the linear or at most quadratic mathematical programming forms, continuous, integer, binary, and mixed, and then solved with professional solvers, using a variety of tricks, which are necessary both in view of the requirements, concerning the nature of solutions to these problems (e.g. binary variables), and the dimensionality of the problems (e.g. the quite frequent very high number of constraints, often in the order of $O(n^v)$, where v may easily exceed 2, and quite often attains 5).

Given the complexity of the thus formulated problems, their computational characteristics, and, what is here most important: the uncertainties, associated with the values of all the coefficients involved (demand, costs, even distances, if one considers farther-off future and the development of the transport networks), it is quite common to refer to the "soft" methods, the ones providing rough approximations of solutions and the possibility of flexible change of the parameters, defining problem shape.

2. The bi-partial approach

The bi-partial approach was developed by the present author at the beginning of the 1980s (see Owsiński, 1980, 1981), primarily as a way of dealing with the general problem of cluster analysis[2]. The strongest point, and, actually, the essence of the bi-partial approach was the capacity of providing the wholesome solution to the clustering problem, including the optimum number of clusters, without the need of referring to any external (usually statistical) criteria. The approach has been recently described in a formal manner in Owsiński (2011, 2012a), and its application to some special narrow task in data analysis was provided in Owsiński (2012b). Then, Dvoenko (2014) applied the approach to extend in this vein the well-known k-means-type procedure.

The approach is based on the construction and use of the *bi-partial objective function*, this function being composed, according to the name, of two terms, which, in a very general way, can be subsumed for clustering as representing, respectively, the *inner cohesion of the clusters* and the *outer separation of the clusters*[3]. Cohesion within clusters is measured by some function of distances between the objects, or measurements, or samples, inside individual clusters, this function being defined over the entire partition of the set of objects, and denoted $Q_D(P)$, where P is a partition of the set of n objects, indexed $i = 1,\ldots,n$, into clusters A_q, $q = 1,\ldots,p$, and subscript D means that we consider distances inside clusters. The counterpart measure of separation of different clusters is denoted $Q^S(P)$, where we mean a function of similarities of objects in different clusters. The sum of the two, denoted $Q_D^S(P)$,

$$Q_D^S(P) = Q_D(P) + Q^S(P),$$

[2] Although, see Owsiński (2011, 2012a), the approach, in a slightly different variant, can be quite effectively used, as well, and that with an interesting generalising interpretation, to deal with the problem of aggregation of orderings (rankings).

[3] In some other circumstances the two aspects, or components, can be referred to as "precision" and "distinguishability", which, in turn, brings us quite close, indeed, to the standard oppositions, known from various domains of data analysis, such as "fit" and "generalisation", or "precision" and "recall".

is minimised, meaning that we seek possibly small distances inside clusters and possibly small similarities among clusters.

This function, $Q_D^S(P)$, has a natural dual, correspondingly denoted $Q_S^D(P)$, in which the two components represent, respectively, cohesion within clusters, measured with similarities (proximities) inside the particular clusters, $Q_S(P)$, and distances between different clusters, measured with distances between objects, belonging to different clusters, $Q^D(P)$. The function $Q_S^D(P)$,

$$Q_S^D(P) = Q_S(P) + Q^D(P),$$

is, of course, maximised.

It should be noted that we do not impose in this general formulation any assumptions nor constraints on the way partitions are conceived (e.g. overlapping, or fuzzy, or rough sets, as clusters), except for the fact that the clusters have to exhaust the entire set of objects considered, just as we do not impose any general form of $Q_D^S(P)$ or $Q_S^D(P)$ and their respective components. Actually, there exist definite conditions that may be set on the components of $Q_D^S(P)$ and $Q_S^D(P)$, but these conditions are meant to secure the capacity of devising an effective solution seeking procedure (see Owsiński, 2012a), and not just the essential rationality of the approach, which is supposed to provide the possibility of representing the generic problem of cluster analysis, and the capacity of comparing the quality of solutions (partitions).

Even though the concept, at its general level, as outlined here, may appear to be close to trivial ("putting similar together and dissimilar apart"), first, it corresponds exactly to the generic formulation of the problem of cluster analysis, and, second, there exist concrete implementations of the two dual objective functions, which form novel and interesting approaches, both regarding the problem of cluster analysis and the one of aggregation of orderings. Moreover, as mentioned above, if the components of the objective functions are endowed with definite, quite plausible properties, the approach leads to effective solution algorithms.

3. The modified bi-partial problem formulation for the facility location

We shall now attempt to form a link between the two domains here introduced, that is – the facility location problem and the bi-partial approach. This link has been already vaguely signalled by mentioning that the facility location problem may turn out to be too complex to be effectively solved by formulating the standard mathematical programming problems and using respective methods and (usually) commercial software, and that, therefore, other kinds of techniques – providing some sorts of approximations – might be referred to. These techniques might, in particular, be taken from the broad domain of cluster analysis. And the bi-partial approach is exactly designed to solve the basic problem of cluster analysis (the application to the aggregation of orderings being, in a way, a side-effect).

It should be emphasised that the problem that we address here is different from the majority of problems, which are considered as instances of application of the bi-partial approach. Namely, the problem we address is, explicitly, a classical question in operations research, related to location analysis. Not only, though, the interpretation of the problem is quite specific, relatively remote from the core of data analysis, but also the very form the problem takes is in a way not appropriate for the treatment through the bi-partial formalism, as introduced here.

We deal, namely, in a very simplistic, but also very general manner, with the following problem

$$\min \Sigma_q(\Sigma_{i\in Aq}\, w_i d(x_i,x^q) + c(q)) \tag{3}$$

with minimisation being performed over the choice of the set of p points (objects) x_i that are selected as the central or median points x^q, $q = 1,\ldots,p$, along with the determination of the catchment subsets A_q, which are assigned to points x^q.

For our further considerations it is of no importance whether the points x^q, to be found, belong to the set X of objects (medians) or not – i.e. they are only required to be the elements of the space E_X (centres), to which all the objects, either actually observed, or potentially existing, belong. It is, however, highly important that the second component of the objective function in (3), namely $\Sigma_q c(q)$, does not involve any notion of distance or proximity, as formally postulated in the bi-partial approach.

Thus, while $d(.,.)$ is some measure of distance, like in the general formulation of the bi-partial approach, where it enters either $Q_D(P)$ or $Q^D(P)$, $c(q)$ is a non-negative value, interpreted as some cost, related to a facility q and to the subset A_q. The problem (3) is interpreted as the one of finding a set of p ($q = 1,\ldots,p$) locations of facilities, such that the overall cost, composed of the sum of distance and weight related costs between points, assigned to the individual facilities (forming subsets A_q), and these facilities, and the sum of costs, related uniquely to these facilities (even though possibly through a function of characteristics of the catchments A_q), is minimised. It is, of course, assumed that the costs $c(q)$ and distances $d(.,.)$ are appropriately (mutually) scaled, in order for the whole to preserve the interpretative sense.

The costs $c(q)$ may be given in a variety of manners: as equal constants for each arbitrary point from X or from E_X, i.e. c, so that the cost component in (3) is simply equal pc, or (more realistically) as the values, determined for each point separately, i.e. $c(i)$, or as a function, composed of the setup component (say, c_1, if this setup cost is equal for all locations) and the component that is proportional to the number of locations, assigned to the facility q, with the proportionality coefficient equal c_2 (that is, the cost for a facility is then equal c_1 + card$A_q c_2$). Of course, more complex, nonlinear cost functions, including also those with c_1 replaced by $c_1(i)$, can, as well, be (and sometimes are) considered. Additional complexity is brought by the consideration of demands, specific for each location, or transport flows (including the maximum admissible flows) between locations. Some of these additional aspects can, though, be relatively easily accommodated within the basic formulation (3).

This problem has a very rich literature, with special numerical interest in its "pure" form, without the cost component, mainly devoted to mathematical and geometric properties and the respective (approximation) algorithms and their effectiveness. Some of the prominent examples of studies in this domain are Hochbaum (1982), Li (2011), Arora, Raghavan and Rao (1998), or Thorup (2001). As mentioned at the beginning, the origin of the practical aspect of the problem is constituted by the so-called Weber problem of location on the plane, which got increasingly complex over time, with a vast array of variants, including the substantive and numerical aspects already mentioned. Notwithstanding this abundant tradition, the issues raised within it and the results obtained, we shall consider here the form of (3) in one of its basic variants.

4. The not so straightforward way to cluster analysis

Any Reader with some basic knowledge in cluster analysis shall immediately recognise the first component of (3) (and, of course, of (1) and (2)) as corresponding to the vast family of the so-called "k-means" algorithms, where such a form is taken as the minimised objective function. Indeed, this fact is the source of numerous studies, linking facility location problems with clustering approaches. One can cite in this context, for instance, the work of Pierre Hansen (e.g. Hansen et al., 2009), or, in a different perspective, that of Furuta et al. (2007), but most to the point here is the proposal from Liao and Guo (2008), this proposal explicitly linking k-means with facility location, similarly as this was done several decades ago by Mulvey and Beck (1984).[4]

The proposal from Liao and Guo (2008) is insofar quite straightforward, but also interesting in practical terms, as the facility of realisation of the basic k-means algorithm allows for the relatively uncomplicated accommodation of additional features of the facility location problem (e.g. definite constraints on facilities and their sets).

Thus, while the first component of the function (3) could be treated with some clustering approaches, e.g. those based on the k-means type of procedure, the issue is in the way the entire function (3) is to be minimised.

At this point it must be indicated, and, indeed, emphasised, that there is no simple analogy between the k-means-like algorithms and the facility location problem, as expressed through (1), (2) or (3). The issue resides in the fact that the 'mean' is, in general, by no means the proper solution to the problem of minimum distance sum, this fact being exactly the origin to the classical Fermat-Torricelli problem[5]. Let us illustrate this fact with an example, which will be used later on to also illustrate some other aspects of this essential question.

Assume, namely, that we deal with a unidimensional situation, and consider three cases as follows:

Case, n =	Positions of points (locations i)	Mean position*	Sum of distances from the mean	Sums of distances from the selected points (i) having the positions x_i:	
				2	3
1, n=4	0, 2, 3, 11	4	14	12	12
2, n=5	0, 2, 3, 5, 15	5	20	18	19
3, n=5	0, 2, 3, 5, 10	4	14	14	13

* for the sake of simplicity of this example, it was assumed that we refer to the mean, and not to the medoid (i.e. the location the closest to the mean among the x_i)

[4] Even a glimpse at the article on the "Facility location problem" in Wikipedia and the list of references therein makes absolutely obvious the close association between this problem and clustering. Yet, the reservations, made explicit here and commented upon, have to be kept in mind.

[5] It is characteristic that many studies, in which clustering is used as a heuristic approach to solve the facility location problem, do not even mention this basic issue, this being, in particular, the case of Liao and Guo (2008), quoted before.

This trivial example shows (and/or implies) that,

(i) indeed, the mean is not the solution to the original Fermat-Torricelli problem, and

(ii) the error, resulting from taking the mean as solution, is brought close to the minimum (ultimately: arbitrarily close to zero), when: (ii.a) the potentially existing 'outliers' are not too distant or nonexistent, (ii.b) the objects form dense clusters, and there are no significant irregularities to the distribution of points i. Indeed, if facilities are located at some points x_i and the density of the points, assigned to the catchments A_q (and, of course, even more so, if the overall density of points in X) is sufficiently high, the error here considered can be altogether neglected in view of the potential errors (or "uncertainties") from other sources, indicated before.

It might be added, concerning the influence of outliers on the difference of results that we here focus on, that if the outliers are really so important – they might be either assigned separate facilities, if their demand (w_i) justifies such a decision, or neglected in the solution, if w_i is low enough[6].

Summing up – clustering can provide an approximation to the solution of the facility location problem, this approximation being insofar justified that the level of uncertainty, involved in the concrete formulations of the facility location problems (especially the more complex ones) is indeed very high.

5. Some other simple examples

In order to develop the simple examples, illustrating both the problem at hand and the possibility of using the bi-partial methodology in solving it, we shall consider the formulation (3) in the following more concrete, even though still very simple, indeed next to trivial, variant:

$$\min_P \Sigma_q(\Sigma_{i\in Aq}\, d(x_i,x^q) + c_1 + c_2\mathrm{card}(A_q)) \quad (4)$$

where c_1 is the (constant) "facility setup cost", while c_2 is the (constant) unit cost, associated with the servicing of each object $i \in A_q$, except for the "first one", this latter cost being included in the setup cost ("once we serve at least one unit of demand"). Such a formulation, even if still quite stylised, seems to be fully plausible as an approximation. It can, of course, be transformed to

$$\min\, (\Sigma_q\Sigma_{i\in Aq}\, d(x_i,x^q) + pc_1 + c_2n), \quad (4a)$$

where it becomes obvious that we could deal away with the component, associated with the unit cost c_2. We shall keep it, though, for illustrative purposes, since the part, related to unit costs may, and usually actually does, take more intricate, nonlinear forms.

The problem (4) can be, quite formally, and with all the obvious reservations, mentioned, anyway, before, moulded into the general bi-partial scheme, i.e.

$$\min_P Q^S{}_D(P) = Q_D(P) + Q^S(P), \quad (5),$$

[6] In terms of the operational research this might be expressed through the element of the objective function that assigns lower "transport capacity" (frequency, volume,...) to the outlying location of a catchment.

where partition P encompasses, in this case, both the composition of A_q, $q = 1,\dots,p$, taken together with the number p of facilities, and the location of these facilities, i.e. choice of locations from (say) X, as the places for facilities q.

Now, consider the simple case, shown in Fig. 1, with $d(.,.)$ defined as Manhattan distance (i.e. the sum of the absolute differences of values along the individual dimensions), the cost component of (4) being based on the following parameter values: $c_1 = 3$, $c_2 = 1$. Again, these numbers, if appropriately interpreted, can be considered plausible (e.g. distance, which corresponds to the annual transport cost, and c_1, which corresponds to annual write-off value).

For this example, Table 1 shows the values of $Q^S{}_D(P) = Q_D(P) + Q^S(P)$, according to (4), for a series of partitions P. We consider in this example a nested set of partitions, i.e. in each consecutive partition, constituting the series, one of the subsets of objects, a cluster A_q, is the sum of some of the clusters from the preceding partition, while all the other clusters are being preserved. Such a nested sequence of partitions is characteristic for a very broad family of cluster algorithms, namely the progressive merger or progressive split algorithms.

The character of results from Table 1, even if close to trivial, is quite telling, and indeed constitutes a repetition of the observations made for multiple other cases, in which the bi-partial approach has been applied, associated with the cluster analysis problems. Note, in particular, that the values of $Q_D(P)$ increase along the series of partitions, while the values of $Q^S(S)$ – decrease, and the sum of the two, $Q^S{}_D(P)$ has a minimum, which, for his simple case, corresponds, indeed, to the solution to the problem.

It is, namely, well known that the optimum (minimum) values of $Q_D(P)$, defined as in (4), that is – as a simile of the classical k-means objective function – decrease with the number of clusters, p, reaching $Q_D(P) = 0$ ("global optimum") for $p = n$, i.e. when each object (observation) is a separate cluster, and the objects are their own (obviously, the best!) representatives. In the case of the function (4), the second component increases with p, also in quite a natural manner. We deal, therefore, with the opposite monotonicity of the two series of values, this opposite monotonicity of the optimum values for the consecutive numbers of clusters, p, being one of the basic conditions for the rationality of the bi-partial objective function.

For a simple, but teling, comparison, Table 2 shows the results for the very same sequence of partitions, but for a somewhat different objective function, (4b), which is formulated namely, as:

$$\min_P \Sigma_q(\Sigma_{i\in Aq}\, d(x_i,x^q) + c_1 + c_2(\mathrm{card}(A_q)\text{-}1)) \tag{4b}$$

or, equivalently,

$$\min_P (\Sigma_q\Sigma_{i\in Aq}\, d(x_i,x^q) + c_1 p + c_2(n\text{-}p)) \tag{4b'}$$

where, however, the setup cost (which, it is implied, as mentioned before in the description of assumptions to this formulation of the problem, contains the cost of the first location served), is $c_1 = 5$, while $c_2 = 1$, as before. Although the minimum appears to be placed in the same configuration, the character of the curve of objective function values changes a bit.

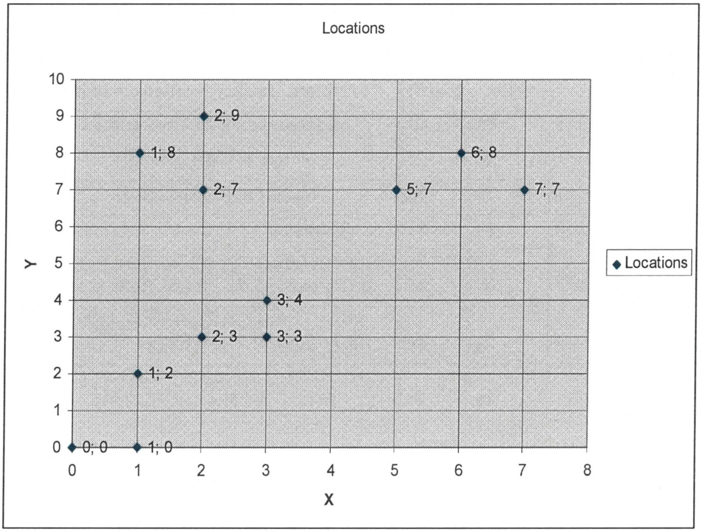

Figure 1. Data for a simple academic example of the facility location problem

Table 1. Values of $Q^S_D(P) = Q_D(P) + Q^S(P)$ for a series of partitions, according to formula (4), with $c_1 = 3$ and $c_2 = 1$

$Q_D(P)$	$Q^S(P)$ – calculation	$Q^S(P)$ - value	$Q^S_D(P)$	Partitions (facility locations in bold)	p
0	12*3+12*1	48	48	All locations are facility locations	12
1	11*3+10*1+1*2	45	46	Merger of (0,0) and (1,0)	11
2	10*3+8*1+2*2	42	44	Merger of (2,3) and (3,3)	10
3	9*3+7*1+2+3	39	42	Addition of (3,4) to (2,3) and (3,3)	9
13	4*3+4*3	24	37	{(0,0), (**1,0**), (1,2)} {(2,3), (**3,3**), (3,4)}, {(5,7), (**6,8**), (7,7)}, {(**1,8**), (2,7), (2,9)}	4
22	3*3+6+3+3	21	43	{(0,0), (1,0), (**1,2**), (2,3), (3,3), (3,4)}, {(5,7), (**6,8**), (7,7)}, {(**1,8**), (2,7), (2,9)}	3
55	1*3+12	15	70	{(0,0), (1,0), (1,2), (2,3), (3,3), (**3,4**), (5,7), (6,8), (7,7), (1,8), (2,7), (2,9)}	1

Table 2. Values of $Q^S_D(P) = Q_D(P) + Q^S(P)$ for a series of partitions, according to (4b), with $c_1 = 5$ and $c_2 = 1$

$Q_D(P)$	$Q^S(P)$ – calculation	$Q^S(P)$ - value	$Q^S_D(P)$	Partitions (facility locations in bold)	p
0	12*5	60	60	All locations are facility locations	12
1	11*5+1*1	56	57	Merger of (0,0) and (1,0)	11
2	10*5+2*1	52	54	Merger of (2,3) and (3,3)	10
3	9*5+1+2	48	51	Addition of (3,4) to (2,3) and (3,3)	9
13	4*5+4*2	28	41	{(0,0), **(1,0)**, (1,2)} {(2,3), **(3,3)**, (3,4)}, {(5,7), **(6,8)**, (7,7)}, {**(1,8)**, (2,7), (2,9)}	4
22	3*5+5+2+2	24	46	{(0,0), (1,0), **(1,2)**, (2,3), (3,3), (3,4)}, {(5,7), **(6,8)**, (7,7)}, {**(1,8)**, (2,7), (2,9)}	3
55	1*5+11	16	71	{(0,0), (1,0), (1,2), (2,3), (3,3), **(3,4)**, (5,7), (6,8), (7,7), (1,8), (2,7), (2,9)}	1

6. Some algorithmic considerations: the reference to the k-means procedure

We have already mentioned that, potentially, the problem naturally lends itself to the k-means-like procedure, due mainly to the form of the first component of the objective function, involving the distances inside the catchments, and that despite the reservations, concerning inappropriateness of the mean as the local solution. Let us remind that the classical k-means procedure, in general and quite rough terms, at that, takes the following course:

0° Generate p[7] points as initial (facility location) seeds (in this case, the case of p-centers, the points generated belong to X), usually $p << n$

1° Assign to the facility location points all the n points (locations) from the set X, based on minimum distance, establishing thereby clusters (catchments) A_q, $q = 1,\ldots,p$

2° If the stop condition is not fulfilled, determine the representatives (here: facility locations) for the clusters A_q, otherwise STOP

3° Go to 1°.

The stop condition is, in a natural manner, constituted by the lack of changes in the assignment of locations to catchments. It may also be a certain minimum degree of changes, or a repetition of the assignments, after a number of iterations. Due to the known feature of fast convergence, the stop condition can also simply be constituted by the predefined limit number of iterations.

Although the procedure converges very quickly, it can get stuck in a local minimum. Yet, owing to its positive numerical properties, it can be restarted from various initial sets of p

[7] We use the classical name of the k-means algorithm, although the number of clusters, referred to in this name as "k", is denoted in the present paper, conform to the notation adopted in the bi-partial approach, by p.

points many times over, and the minimum values of the objective function obtained indicate the proper solution (usually, a vast majority of results corresponds to the same configuration, while the minority are the non-optimal local minima).

In the here analysed problem of facility location, since such problems rarely are really large in the standard sense of data analysis problems, it is indeed quite feasible to run the k-means procedure, as outlined above, for the consecutive values of p in order to check whether a minimum over p can be found for a definite formulation of the facility-location-related $Q^S_D(P)$. Although we shall not be demonstrating this here, let us note that in view of the opposite monotonicity of the two components of $Q^S_D(P)$ along p, the minimum found over p is a global minimum (although, of course, it is not necessarily the solution to the problem considered, since we deal here only with an approximation of the actual proper objective function). This procedure can be simplified so as to encompass only a part of the sequence of values of p, starting, say from $p = 2$ upwards, until a (single) minimum is encountered.

7. Algorithmic considerations based on the bi-partial approach

The direct application of the bi-partial algorithmic precepts

We shall now present the algorithmic approach that is founded on the basic precepts of the bi-partial approach. Assuming, namely, the property that we have observed for the case of the here considered concrete objective function (4), that is – the opposite monotonicity of the two components of the objective function, we can reformulate it, obtaining, in the general case, the following parametric problem:

$$\min_P Q^S_D(P,r) = rQ_D(P) + (1\text{-r})Q^S(P), \tag{6}$$

where the parameter $r\in[0,1]$ corresponds to the weights we might attach to the two components of the objective function. Actually, this parameter is used exclusively for the algorithmic purposes, and it is not meant to express any sort of substantive weight, as we assume that we ultimately weigh equally the two components (i.e. $r = ½$). Here, we make no a priori assumptions as to the value of p, in distinction from the approach, outlined above, based on the classical k-means procedure. The form (6) enables the construction of a suboptimisation algorithm, provided the two components of the objective function are endowed with certain properties. We shall outline here the construction of this algorithm for the case of the objective function (4), for which at least an important part of the relevant properties is satisfied (see Owsiński, 2011, for a more complete account on the conditions for the algorithmic effectiveness of the bi-partial approach).

Thus, the above general form is equivalent, for (4), to the following one:

$$\min_P (r\Sigma_q\Sigma_{i\in Aq}\, d(x_i,x^q) + (1\text{-}r)\Sigma_q(c_1 + c_2\text{card}A_q)). \tag{7}$$

Now, take the iteration step index, t, starting with $t = 0$, and calculate the successive values of r^t, corresponding to the optima of (7) for the consecutive steps of the proposed procedure. Consider (7) for $r^0 = 1$. We obtain

$$\min_P (1\cdot\Sigma_q\Sigma_{i\in Aq}\, d(x_i,x^q) + 0\cdot\Sigma_q(c_1 + c_2\text{card}A_q) = \Sigma_q\Sigma_{i\in Aq}\, d(x_i,x^q)). \tag{8}$$

Since we did not make any assumptions, concerning the value of p, we can easily see that the global minimum for (8) is obtained for $p = n$, i.e. when each object (location) contains a facility (that is: each location constitutes a separate cluster). Denote this particular, extreme (but also, in a way, “optimal”) partition by P^0. The situation described is illustrated in the first line of Table 1. The value of the original objective function is, therefore, equal $n(c_1 + c_2)$, since the first component disappears, we deal with n facilities, and all cardA_q = cardA_i are equal 1.

Then, we decrease the value of r from $r^0 = 1$ down. At some point, for the value of the parameter r that we denote r^1, this parameter value is low enough to make the value of the second component of the objective function, $(1\text{-}r)\Sigma_q(c_1 + c_2\text{card}A_q)$, weigh sufficiently to warrant aggregation of two locations into one cluster, meaning that one facility would serve the two locations aggregated at this step. This happens when the following equality holds:

$$Q^S{}_D(P^0,r^1) = Q^S{}_D(P^1,r^1), \tag{9}$$

where P^1 denotes the partition, which corresponds to the result of the initial aggregation operation mentioned, the equality from (9) being equivalent, in the case here considered, to

$$r^1\cdot 0 + (1\text{-}r^1)\, n(c_1 + c_2) = r^1 d(i^*,j^*) + (1\text{-}r^1)\,(n(c_1 + c_2) - c_1) \tag{10}$$

where i^*,j^* is the (aggregated) pair of locations, for which the value of r^1 is determined. This value, conform to (10) equals

$$r^1(i^*,j^*) = c_1/(d(i^*,j^*) + c_1). \tag{11}$$

The simple relation thereby obtained, which determines the rule for aggregating, at least, the pairs of particular locations into clusters (the case of aggregation of larger components being left out for a while) is justified by the fact that for each passage from some p to p-1, accompanying such an aggregation, the value of the second component decreases by c_1, while a value of distance, or a more complex function of distances, is added to the first component.

Since we look for the highest possible r^1, which follows $r^0 = 1$, it is obvious, also from (10) and (11), that the $d(i^*,j^*)$ we are looking for must be the smallest one among those not yet contained inside the clusters (i.e., for this initial aggregation step – the smallest one among all the distances between objects). In the subsequent steps t we shall be using the equation (9) in its more general form, referring to the general formulation of the objective function, i.e.

$$Q^S{}_D(P^{t-1},r^t) = Q^S{}_D(P^t,r^t), \tag{12}$$

and derive from it the expression analogous to (11). In this particular case – which is, anyway, quite similar to several of the implementations of the bi-partial approach for clustering – the equation, analogous to (11) is obtained from (12), meaning that at each step t the minimum of distance is being sought (although, in general, not just between the individual locations), exactly as in the classical progressive merger procedures, like single link, complete link etc.

The procedure stops when, for the first time, r^t is obtained in the decreasing sequence of $r^0, r^1, r^2,\ldots$, having the value lower than ½ (the sequence of r^t, if realised until the aggregation of all locations into one cluster, will, of course, end at $t = n$-1). Falling below ½ means, namely, that “on the way” the partition P^t was obtained, which was generated by the algorithm for r = ½, corresponding to the equal weights of the two components of the objective function.

Thus, on the one hand, we deal with a procedure that is entirely analogous to the very popular, simple progressive merger algorithms, but, on the other hand, has an inherent capacity of indicating the "solution" to the problem, without any reference to an external criterion, rather than just generating the all-embracing dendrogram. We used the quotation marks, when speaking of "solution", because the procedure does not guarantee in any way the actual minimum of (5) (or the particular specific embodiments of this general function, like, in this case, (4), with its variants), since the operations, performed at each step, are limited to aggregation. The experience with other cases shows that a simple search in the neighbourhood of the suboptimal solution found suffices for finding the actual solution, if it actually differs from the suboptimal one.

It is left to the Reader to see that if we use the definition (4b) instead of (4), the procedure derived following the above pattern is exactly the same, even though the concrete algebraic formulae slightly differ.

In general, the logic consists in the construction of a simile of (6), which, provided the involved components of the objective function behave not only so that they display opposite monotonicity (this being the necessary and quite obvious condition, which is fulfilled by virtually all "reasonable" formulations), but also that this monotonicity takes place appropriately along the branches of the dendrogram, secures the possibility of implementing the simple sub-optimisation procedure here outlined. The potential improvements over such a procedure are left to the considerations, associated with the particular cases, for which the approach is being implemented.

It is obvious that the procedure here proposed is not only an approximation from the point of the optimum to the facility location problem, but also – with respect to the very clustering procedure, since we admit only aggregating operations. There may exist cases, when such a procedure does not yield the optimum solution. Yet, for "appropriately well behaved" data, meaning lack of cumbersome shapes of clusters, first of all, this approximation works relatively well.

The algorithmic return to the k-means paradigm

It is, however, also possible to link the two paradigms, that is – the one of the bi-partial approach and of the k-means procedure, here – in the context of the facility location problem. Let us return, namely, to the formulation (4b):

$$\min_p \left(\Sigma_q \Sigma_{i \in A_q} d(x_i, x^q) + c_1 p + c_2(n-p)\right) \qquad (4b')$$

and assume, quite reasonably, that $c_1 > c_2$ (or even $c_1 >> c_2$). Under such circumstances the procedure, which refers to the k-means paradigm, would be as follows:

-- find the k-means-type solution to the clustering problem with $d(x_i, x^q)$ for a sequence of values of p, potentially starting from $p = 1$, thus obtaining the (approximation to the) minimum values of the internal sum of distances (recalling here that the approximation results from the fact that the k-means minimise the sums of squares of distances involved), i.e.

$$\Sigma_q \Sigma_{i \in A_q} d(x_i, x^q)$$

that we can denote $D_P^*(p)$, $p = 1, 2,\ldots$; this sequence is decreasing, of course;

-- calculate the values $D_P^*(p) + p(c_1 - c_2)$, the second component neglecting the value $c_2 n$ from (4b') as constant, this second component constituting a sequence increasing linearly with p;

-- we look for the minima in the thus obtained sequence of values; even though there may theoretically exist more than one local minimum along p (see further on for a comment), it should suffice to stop at the very first such minimum.

An illustration, based on the data for the previously presented trivial example, is provided in Fig. 2 (the data, corresponding to this diagram, are provided in Table 3). The respective minimum is not only singular, but is also very clearly seen. Let us once again emphasise that we deal here with definite approximations: (a) k-means approximating the centre sought; (b) the choice of the minimum, if there may be more of them, and, of course, (c) the entire formulation, approximating the true economic aspect of the problem. With respect to point (b) the essential aspect is associated with the choice of the locations – main influence being associated with the assignment of locations to facilities from among i. all of the potential locations, ii. all of the existing ("demand") locations, iii. only a subset of selected locations. Then, regarding point (c), it is of foremost importance to secure monotonicity of the facility cost component of the objective function.

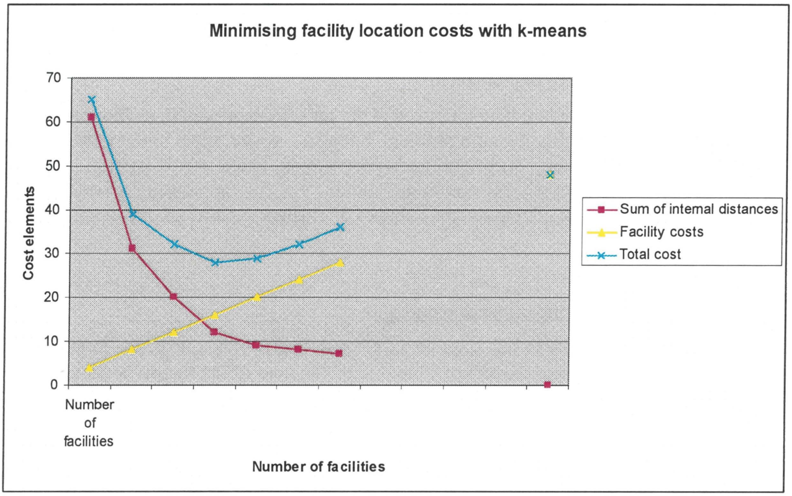

Figure 2. An example of application of the bi-partial concept, based on the k-means-like procedure (data from Fig. 1)

With respect to the trivial experiment, whose results are shown in Fig. 2 and Table 3, let us add, in the context of the above remarks, that it was assumed that the facility locations can not only belong to the set of the "demand locations" (having equal weights – "demands"), but can also be picked at the integer nodes of the coordinate grid on the plane. The assumption of choosing only the "demand locations" as potential facility locations would have (the medoid algorithm), at least theoretically, shifted the position of the minimum with respect to p towards some lower value, since the values of distance sums, obtained from the k-means type process would tend to be higher at the beginning, i.e. would not decrease that rapidly along p.

Table 3. Account on the experiment, illustrated in Fig. 2, with data originating from Fig. 1

Number of facilities, $p =$	Locations of facilities on the plane, (x,y)	Division of catchments among facilities	Values of objective function components and its total		
			$D_P^*(p)$	$p(c_1-c_2)$	$D_P^*(p)+p(c_1-c_2)$
1	4,5	all locations in one group	61	4	65
2	2,2; 4,8	division into lower and upper groups	31	8	39
3	2,2; 2,8; 6,7	upper group is split	20	12	32
4	**1,1; 3,3; 2,8; 6,7**	**lower group is split**	**12**	**16**	**28**
5	1,0; 1,2; 3,3; 2,8; 6,7	lower group is further split	9	20	29
6	1,0; 1,2; 3,3; 1,8; 2,8; 6,7	left upper group is split	8	24	32
7	1,0; 1,2; 3,3; 1,8; 2,8; 6,7; 6,8	right upper group is split	7	28	35
12	facilities located at all demand locations		0	48	48

$c_1 = 5$, $c_2 = 1$

Let us also note that the above proposed procedure may be, actually, used with any kind of the k-means-like algorithm, including the fuzzy-set-based ones (for the initial work on this family of methods, see Dunn, 1974), like the classical FCM. In that case, the already mentioned managerial issue arises of the significance of the fact that particular demand locations i are assigned to the facilities q to a certain degree, $\mu_{iq} \in [0,1]$ through minimisation of the objective function

$$\Sigma_q \Sigma_i \mu_{iq}^{\alpha} d^2(x_i, x^q),$$

leading to the determination of the $n \times p$ matrix of the memberships μ_{iq}, where, additionally, the exponent $\alpha > 1$ establishes the "degree of fuzziness" of the clusters (facility catchments) obtained from the FCM procedure. For $\alpha = 1$ one obtains from this procedure the "crisp" clusters, corresponding to those, resulting from the classical k-means (see Bezdek and Pal, 1992), while for increasing values of α the μ_{iq} tend to a uniform distribution, in the limit all being equal $1/p$. Thereby, an instrument is provided for controlling the resulting values of μ_{iq}, so that for "reasonable" values of the exponent α only few of the memberships μ_{iq} differ from 0 or 1. Like in the previously indicated case of cliques, these may be indicatively used as "emergency options", with an indication of preference being established through the value of μ_{iq}.

7. Some comments and the outlook

Extensions

The illustration, here provided, even though extremely simple, is definitely sufficient to highlight the capacity of the bi-partial approach to deal with the p-median / p-center type of facility location problems. In fact, for (slightly) more complex formulations of the problem, like, say

$$\min_P \Sigma_q(\Sigma_{i\in Aq}\, d(x_i,x^q) + c_1(q) + c_2 f(\mathrm{card}(A_q))) \quad (13)$$

i.e. where setup costs are calculated for each potential facility location separately, and $f(.)$ is an increasing concave function, the relation analogous to (12) yields only marginally more intricate procedure, analogous to that based on (11), where for each aggregation the minimum has to be found for the two locations or clusters aggregated.

The issue, worth investigation, which arises therefrom is: what realistic class of the facility location problems can be dealt with through the bi-partial approach?

Another case that is of interest, indeed, in the context of facility location, exceeding the capacities of the standard approaches, based on mathematical programming, is associated with the quite natural formulation of the facility location problem, in which a hierarchy of centres is designed, where facilities are located, featuring varying properties, depending upon the level of the hierarchy.

For this case, let us introduce a slightly more complex notation. Thus, let h denote the level of hierarchy, with $p^h(q^h)$ being the number of catchments, determined for the level h and the catchment (of the preceding level) number q^h. The numbering of levels starts here with $h = 1$ for the catchment $q^0 = 1$, which encompasses the entire set of locations, $x_1,\ldots,x_n$. The respective minimisation problem, a simile of what we have been considering, might have the form of

$$\min_{P^H} \sum_{h=H}^{h=1} \sum_{q^h=1}^{p^h} (\sum\nolimits_{i\in A_{q^h}} d(x_i, x^{q^h}) + c_1(q^h) + c_2 f(card(A_{q^h}))) \quad (14)$$

where P^H denotes the hierarchy of partitions into catchments A_{q^h}, with the partitions at levels h being constituted by the sets of catchments $\{A_{q^h}\}_{q^h=1}^{q^h=p^h}$. Once the data for such a hierarchical facility location problem are given, the procedure, analogous to the one here outlined, might be used, with additional assumptions, concerning the working of the procedure at the consecutive levels h.

Some conclusions

The actual design of the procedure, based, on the one hand, on the precepts of the bi-partial approach, and on the other hand – on the principles of the k-means-type algorithms, shall depend upon the shape of the problem at hand. In this context, the following remarks might be forwarded:

(1) k-means outperform progressive merger procedures for data sets with numerous objects (locations), but not too many dimensions (here: by virtue of definition, either very few, or just two), when storing of the distance matrix and operating on it is heavier than calculating np (much less than n^2) distances at each iteration; in the cases envisaged n would not exceed thousands, and p is expected not to be higher than 100, so that the two types of procedures might be quite comparable;

(2) there exists a possibility of constructing a hybrid procedure, in which k-means would be performed for a sequence of values of p at the later stages of the bi-

partial procedure, with the result of the aggregation, performed by the bi-partial procedure being the starting point for the k-means algorithm;

(3) given the proposal by Dvoenko (2014), there exists also a possibility of implementing directly the bi-partial version of k-means, with specially designed form of the two components of the objective function; this, however, would require, indeed, additional studies.

References

Arora, S., Raghavan, P. and Rao, S.: Approximation schemes for Euclidean *k*-medians and related problems. *STOC'98. Proc. of the 30th annual ACM symposium on theory of computing.* ACM, New York, 1998.

Bezdek, J. C., Pal, S. K.: *Fuzzy Models for Pattern Recognition. Methods that Search for Structures in Data.* IEEE, New York, 1992.

Dunn, J. C.: A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *J. Cyber.* **3**(3), 32-57, 1974.

Dvoenko, S.: Meanless k-means as k-meanless clustering with the bi-partial approach. Proceedings of PRIP 2014 Conference. Minsk, May 2014.

Furuta, T., Sasaki, M., Ishizaki, F., Suzuki, A. and Miyazawa, H.: A New Clustering Algorithm Using Facility Location Theory and Wireless Sensor Networks. Technical Report of the Nanzan Academic Society. Mathematical Sciences and Information Engineering. NANZAN-TR-2006-04, 2007.

Hansen, P., Brimberg, J., Urosević, D., Mladenović, N.: Solving large p-median clustering problems by primal-dual variable neighbourhood search. *Data Mining and Knowledge Discovery*, 19: 351-375, 2009.

Hochbaum, D. S.: Heuristics for the fixed cost median problem. *Mathematical Programming* **22**, 148-162, 1982.

Li, S.: A 1.488 Approximation Algorithm for the Uncapacitated Facility Location Problem. *Automata, Languages and Programming.* Lecture Notes in Computer Science **6756**, 77-88, 2011.

Liao, K., Guo, D.: A clustering-based approach to the capacitated facility location problem. *Transactions in GIS*, 12(3): 323-339, 2008.

Mulvey, J. M., Beck, M. P.: Solving capacitated clustering problems. *European Journal of Operational Research* 18:339-348, 1984.

Owsiński J.W.: *Regionalization revisited: an explicit optimization approach.* CP-80-26. IIASA, Laxenburg 1980.

Owsiński J.W.: Intuition vs. formalization: local and global criteria of grouping. *Control and Cybernetics*, 10, 1981, no. 1-2, 73-88.

Owsiński, J.W.: The bi-partial approach in clustering and ordering: the model and the algorithms. *Statistica & Applicazioni*, 2011, Special Issue, pp. 43-59.

Owsiński, J. W.: Clustering and ordering via the bi-partial approach: the rationale, the model and some algorithmic considerations. In: J. Pociecha & Reinhold Decker, eds., *Data Analysis Methods and its Applications*. Wydawnictwo C.H. Beck, Warszawa, 2012a, pp. 109-124.

Owsiński J. W., On the optimal division of an empirical distribution (and some related problems). *Przegląd Statystyczny*, special issue 1, 2012b, 109-122.

Owsiński, J. W., Bi-partial version of the p-median / p-center facility location problem and some algorithmic considerations. *JAMRIS*, 2014, 3, 59-63. DOI 10.14313/JAMRIS_3-2014/28

Thorup, M.: Quick *k*-Median, *k*-Center, and Facility Location for Sparse Graphs. In: F. Orejas, P. G. Spirakis and J. van Leeuwen, eds.: *ICALP 2001*. Lecture Notes in Computer Science **2076**, 249-260, 2001.