Redefinition of the Greek electoral districts through the application of a region-building algorithm

Photis, Yorgos N.

University of Thessaly, Department of Planning Regional Development

October 2012

Online at https://mpra.ub.uni-muenchen.de/42398/
MPRA Paper No. 42398, posted 05 Nov 2012 23:29 UTC
REDEFINITION OF THE GREEK ELECTORAL DISTRICTS THROUGH THE APPLICATION OF A REGION-BUILDING ALGORITHM

Yorgos N. PHOTIS
University of Thessaly, Department of Planning and Regional Development, Pedion Areos, 38334, Volos, Greece.

Abstract:
The main purpose of this paper is the formulation of a methodological approach for the definition of homogenous spatial clusters, taking into account both geographical and descriptive characteristics. The proposed methodology, is substantiated by SPiRAL (SPatial Integration and Redistricting ALgorithm), a constrained-based spatial clustering algorithm, whose successive steps focus on the analysis of the characteristics of the areas being integrated, the designation of the spatial clusters and the validity of a joining criterion. We applied the methodological approach and used SPiRAL to solve a realistic electoral redistricting problem. Namely, the redefinition of the electoral districts of the Prefecture of Lakonia in Greece. The results demonstrate an improved layout of the study area’s electoral map as far as the problem’s criteria and constraints are concerned (adjacency, population and size), justifying in this respect the perspectives and potential of our approach in the analysis and confrontation of similar problems.

Keywords: Spatial clustering, GIS, constraint-based algorithm, electoral districts, Greece.

1. INTRODUCTION

The organization and arrangement of space as it is expressed through administrative, economic and political decisions has critical consequences both at a community and individual level. Geographers, planners and policy makers, during such problem solving processes mainly utilise geo-referenced datasets at a regional or urban scale. The central issue in all of the above cases is the designation of spatially contiguous and robust clusters of areas, which should reflect efficient and competitive regions. Consequently the structure and profile of the newly formed spatial units should stem from a well-structured process complying with predefined criteria and constraints.

The problem of partitioning a territory into districts is widely recognized as a difficult multicriteria, combinatorial optimization problem (Bozkaya et al, 2011). In this framework, the advances in technology and the fast emerging field of Geoinformation have provided considerable impetus for new methodological approaches. Especially in cases, where the geospatial nature of the problem is clearly formulating the solution process, the integration of spatial analysis methods and Geographic Information Systems provides increased the capabilities in data editing, analysis and representation that are needed.
Taking into account the importance of the rational arrangement of space, the aim of this paper is to present a comprehensive methodological approach of spatial clustering. More specifically, when neighboring regions merge to form spatial clusters, whose main characteristics are homogeneity and spatial cohesion, the proposed method deals with both proximity and contingency constraints according to a set of predefined criteria. Underlining the importance of adjacency in the definition of spatial clusters the criteria utilized in the process of redistricting mainly refer to geographical characteristics of regions. The method’s successive steps formulate the SPiRAL algorithm (SPatial Integration and Redistricting ALgorithm), which is also described in this paper. The proposed methodology is being applied and tested in the definition of a Greek prefecture’s new electoral districts. The criterion set was the homogeneity of population size corresponding to the number of seats.

The paper is organized in five sections through which, the theoretical background and the methodological framework of the approach are presented. The following section introduces a formal definition of spatial clustering problems and a brief overview of methods and techniques applied when dealing with them. The third section focuses on the proposed methodology and the constrained-based algorithm, which during the fourth section are applied to the reorganization of Lakonia’s electoral districts. The last section contains some concluding remarks mainly dealing with the performance and the effectiveness of both the proposed methodology and the SPiRAL algorithm.

2. SPATIAL CLUSTERING

Clustering is a key issue in spatial data acquisition and mining since its main purpose is to identify subsets of data with similar characteristics (Grekousis et al, 2012). According to Estivill - Castro, Lee and Murray (2001), “Spatial clustering consists of a partitioning set $P = \{p_1, p_2, \ldots, p_n\}$ of geo-referenced point-data in a two-dimensional study region $R$, into homogenous sub-sets due to spatial proximity.” In this respect, the final definition of spatial clusters is largely depended on the parameters, which specify their number and characteristics such as size, distribution, arrangement and dispersion. During the last two decades, the rapid development and widespread implementation of GIS have triggered a substantive increase in problems dealing with the formulation of spatial clusters with high levels of homogeneity as their main goal.

One of the major aspects when dealing with spatial planning problems is that the volume of data to be analysed is, usually, large. Spatial data refer to two-dimensional or three-dimensional points, lines and polygons. Exploratory spatial data analysis (ESDA) provides methods and techniques for assisting the processing and exploitation of spatial information, organised in a complex GIS context (Murray et al., 2001). Some of the most applied methods are those dealing with the spatial distribution and organization of data. Bailey and Gatrell (1995) refer to a set of areal proximity measures, such as the distance of centroids, the common border of polygons or the concept of the spatial moving average in order to create a proximity grid. In the unusual case that the distribution of data forms a grid then the median polish technique is applied in order to clarify the spatial trends and tendencies of data. Even though these techniques have been extensively used in geographical research, they cannot by themselves resolve such complex problems. They have to be adopted in a combinatorial framework in order to effectively analyse and explore spatial information according to planning guidelines and objectives.

On the other hand, analytical methods of spatial clustering can be extremely helpful by providing information about the spatial relationships among data (Duque et al, 2007). They aim at the redistricting of space and the best possible allocation of patterns (Openshaw,
In this respect, the interest is focused on finding ways to divide regions in spatial units characterised by similar features.

In recent research and literature a plethora of spatial clustering procedures have been introduced. The most commonly used are spatial clustering algorithms and the main reason for this is their computational accuracy, speed of calculations and robustness of results. According to Sheikholeslami, Chatterjee and Zhang (2000) they can be divided into four broad categories: partitioning algorithms that optimize an equation, hierarchical algorithms which decompose the database, density-based algorithms which exploit a density equation for the determination of the clusters and grid-based algorithms which reduce the dimensions of the clustering space. Alternative point-oriented approaches are provided by the central point method which exploits the distance of artificial points, the median method which is based on the distances throughout the entire point dataset and the based on triplets method which uses the common node of three polygons in order to identify clusters in space (Murray, 1999; Gebhardt, 2000).

As expected, each method is characterized by advantages and disadvantages and thus their exploitation eligibility depends not only on each decision maker’s preferences, but also on the study region’s geographical profile and attributes. However, in most of the above cases computational demand is extremely high and such methods become time-consuming and burdensome. This critical drawback can be dealt with through the utilisation of Geographic Information Systems whose advanced capabilities in spatial data processing and visualization by maps, graphs, or tables can substantially contribute to their performance and efficiency.

3. METHODOLOGICAL FRAMEWORK

The main aim of the paper is to formulate an alternative and reliable methodology for the formulation of spatial clusters. The proposed methodology should be and is substantiated through a constrained-based algorithm used to organise and regionalize the study area’s spatial units in an objective and systematic manner. In such a framework, the analytical fragmentation of the proposed problem solving procedure formulates a constraint-based region-building algorithm. The methodological framework, which confronts the above problem definition, is described in the following paragraphs and depicted in Figure 1.

![Figure 1. The proposed methodological framework.](image)

Firstly, the problem is defined by its decision variables and constraints determination. The spatial clustering process should succeed in the formulation of homogeneous sub-regions.
of the study area, which would reflect certain descriptive and geographical characteristics. Since the finally defined regions should preserve spatial continuity and exhibit compact geometry as spatial clusters, adjacency of their structural components is the most important parameter. During the second stage of the process data collection and processing take place. Due to the fact that the central issue in such spatial processes is the topological characteristics of each entity, GIS functions and operations are utilized and problem parameters emerge from their geographical features.

Consequently, the spatial attributes of the polygons such as perimeter, area, metacenter (spatial mean), relative position in space and topological information (i.e. right polygon – left polygon) will be exploited to determine adjacency. Having completed the processing of spatial data, our interest is then focused on the joining method of polygons, in order to schematize the predefined number of clusters. The proposed method should be capable of determining the way that the initial cluster centers are proposed, the geographical variables that will be used for clustering, the distribution pattern that should be followed and also the joining criterion. Such a method should be in an algorithmic form paired with a terminating threshold, whose sequential completion of steps will produce continuous spatial clusters.

In the forth phase of the methodological framework and after the implementation of the algorithm results are examined and evaluated. This stage is critical since new data and modified parameters can be fed to and processed by the algorithm, in order to improve the solution performance and the effectiveness of the approach.

Finally, conclusions are derived concerning quantitative metrics of resulting configurations as well as evaluative indicators of the methodology’s integrity, spatial redistricting potential and capabilities.

3.1 The Spatial Integration and Redistricting Algorithm (SPiRAL)

As mentioned above, SPiRAL, the proposed constraint-based algorithm combines proximity and contingency measures. In each of its steps the main consideration is formulation of clusters taking into account the spatial distribution of quantitative data, describing the regions (polygons) to be joined. Moreover, the proposed algorithm SPiRAL is consisted of steps and calculations aiming to assist decision makers in clarifying whether the desirable objective is achieved and met. More specifically, SPiRAL evolves as follows:

**STEP 1: Definition of the ‘join criterion’**

The value of the characteristic that the schematized clusters should comply to, i.e. the join criterion, is determined. The join criterion J, can be a demographic, an economic or a social parameter and its deviation margins are given.

**STEP 2: Determination of the n centroids**

A - If polygons exist that satisfy the join criterion J, they are by default are set as cluster centres.

B - If no polygons satisfy the join criterion J, then the following calculations are performed:

1. The maximum \((X_{\text{max}}, Y_{\text{max}})\) and the minimum \((X_{\text{min}}, Y_{\text{min}})\) value of the spatial average of the remaining polygons is computed.
2. The grid where the \(n\) centers will be located is defined by calculating each quadrat’s side as equal to the difference of the maximum and minimum value of the spatial average for every axe divided by \(n+1\) and adding the result (\(s_1\) for axe \(X\) and \(s_2\) for axe \(Y\)) to \(X_{\text{min}}\) and \(Y_{\text{min}}\) respectively, times the center’s serial number. In this respect, the
coordinates of the first centre will be \((X_{\text{min}}+s_1, Y_{\text{min}}+s_2)\), of the second’s \((X_{\text{min}}+2s_1, Y_{\text{min}}+2s_2)\) etc.

3. Distances of each polygon’s centroid to every grid point are calculated. The centroid with the minimum distance becomes a center, and polygons satisfying the joining criterion are assigned to the closest grid point, which is then excluded from the set.

4. If adjacent polygons - centers exist, then the polygon with the smaller centroid distance from the corresponding point of the grid, becomes a center and all others are excluded. Subsequently, the second closest polygon is selected.

**STEP 3: Formulation of the initial clusters**

Neighboring polygons to each center unit (sharing one or more common borders) are assigned to it.

**STEP 4: Allocation of polygons to the initial clusters**

A - Polygons adjacent to one and only one center are merged to form a cluster.

B - Polygons adjacent to two or more centers are called meso-polygons and are associated with a cluster following a modified process:

a. If a cluster meets the joining criterion after the annexation of the remaining neighboring polygons, then the meso-polygon will be assigned to the cluster which does not.

b. For clusters falling short of the criterion, an Index of Unification \(I\) is calculated, which is a standardized form of the distance between two polygon centroids, according to the following equation:

\[
I = \frac{d_{jk}}{l_{jc}}
\]  \[1\]

where

- \(d_{jk}\) = distance between the centers of two polygons \(j\) and \(k\) and
- \(l_{jc}\) = percentage of the common border between a polygon \(j\) and a cluster \(c\) adjacent to its perimeter.

The index \(I\) is calculated for every neighboring to the cluster polygon.

c. The meso-polygon is allocated to the cluster exhibiting the minimum \(I\) value.

**STEP 5: Threshold criterion**

A - Assigned polygons are no further considered

B - Spatial allocation of remaining polygons continues for each cluster until the join criterion \(J\) is satisfied.

**STEP 6: Allocation of remaining polygons**

After the process’s completion and the definition of clusters that meet the join criterion \(J\), every unassigned polygon is allocated to its first order neighboring cluster. If a polygon is adjacent to more than one clusters, the Index of Unification \(I\) is utilised (Step 4: Bb).

**STEP 7: Evaluation of the clusters homogeneity according to the join criterion**

The standard deviation (\(\sigma\)) for the criterion value is calculated for each cluster according to Equation 2 (Papadimas and Kollias, 1998):
\[ \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2} \quad [2] \]

where
- \( i = 1, \ldots, n \)
- \( x_i \) = the criterion value for each cluster \( i \)
- \( N \) = the number of observations which equals the number of clusters, and
- \( \mu \) = the arithmetic mean of the criterion value for all clusters.

**STEP 8: Enhancement of each cluster’s quantitative characteristics**

A - For clusters not meeting the criterion value, all adjacent polygons are selected.
B - The polygon with the maximum common border percentage with a cluster is assigned to it.
C - Assignment becomes permanent if and only if it improves the standard deviation value with the maximum reduction.
D - If no modifications to the initial polygon set are performed, remaining polygons are processed, according to a descending common border percentage order.
E - Each polygon’s adjusted criterion value for each cluster is calculated in order to identify falling short clusters.
F - Adjacent polygons of remaining clusters are located as follows:
   a. For a remaining falling short cluster, an adjacent polygon is a second order common border polygon, if its first order one is already assigned.
   b. For each redefined falling short cluster, an adjacent polygon is one sharing the maximum percentage of common border with it.
G - Any polygon’s n-order assignment should not split its n-1-order cluster.
H - If a cluster exists, which can be separated from a polygon improving its standard deviation, then the polygon with which shares the minimum common border percentage is moved.
I - Steps A to F are repeated until every cluster in the study area meets the joining criterion J.
J - Algorithm stops when no further replacement improves the standard deviation value.

The above steps describe the analytical sequence of the spatial clustering process of polygons according to the proposed constrained-based algorithm. The methodology can be applied to a plethora of redistricting problems related, for example, to the delineation of administrative regions, electoral districts as well as educational, emergency response and metropolitan service areas. In the following section of the paper, SPiRAL is utilised in a example for an administrative region’s fictitious reorganisation while on the same time, its efficiency is tested and evaluated.

**4. REDEFINITION OF LAKONIA’S ELECTORAL DISTRICTS, GREECE.**

Greek electoral districts, not surprisingly characterised by geodemographic and socioeconomic disparities, also seem to be defined by random boundaries that form unadjusted shapes (Valasaki and Photis, 2005). In this respect, they constitute an intriguing and challenging study area to which, the proposed methodological frameworks as well as SPiRAL were applied in order to redesign one of its existing electoral districts namely, the
Prefecture of Lakonia and formulate new homogenous single-seat units by clustering contiguous municipalities.

4.1 Study Area and Data

The current electoral districts of Greece were established by the unification of a number of municipalities and communes (declared as spatial units after the implementation of law 2539/97 for the ‘Reformation of the 1st Degree of the Local Administration’). Their borders are consistent to the latter and the number of seats varies with relevance to their population size.

Consequently and with respect to problem formulation, structural units are municipalities and communes, whereas reference unit is the electoral district. All data utilized throughout the solving process are geo-referenced, mainly concern geographical and descriptive characteristics of polygons and were provided by the Laboratory for Spatial Analysis GIS and Thematic Cartography at the Department of Planning and Regional Development in the University of Thessaly. The join criterion is defined as the equal distribution of the district’s population per seat and the respective demographic data relate to the 2001 Census of the National Statistics Service of Greece.

Regarding processing stages and environments, the proposed methodology was basically realised in ESRI’s ArcGIS platform. Polygons centroids were selected and the topological data of lines and polygons of the region were calculated. Editing also includes calculations concerning the line percentage of common border of every polygon’s perimeter. In a similar manner, contingency tables were created for the entire set of structural polygons where the adjacency percentage of common border with their first order neighbours was added as an attribute.

4.2 Application

According to the successive steps of SpiRAL, its application starts by determining the join criterion to which spatial clusters should comply. To this end, with the 2001 Census stating that the total population of Greece is 10,259,900, with nearly two thirds living in urban areas and since the number of parliament seats is 288, the join criterion, representing the electoral standard of the country, is 35,624 people. The acceptable deviation, in this case, is set to 30%, therefore the population of any one-seated electoral district may vary from 24,936 to 46,312 people. If necessary, different deviation margins can be set by the decision makers.

Lakonia’s electoral district is located in South Continental Greece, with a population 95,696 people and three (3) respective parliament seats. First, the centres of the initial clusters are defined according to the polygons spatial features. Populations of the district’s municipalities show that there is no structural unit satisfying the join criterion. Consequently, centres are defined with respect to each polygon’s centroid (Step 2B). Taking into account the restrictions that the algorithm sets concerning adjacency, as initial centres are set the municipalities of Githio, Skalas and Niaton (Figure 2).

In the next stage, the remaining polygons are initially allocated to the three centers according to Steps 3 and 4 of the algorithm and simultaneously processed for all districts. More specifically, Anatoliki Mani, Oitilou and Smynoun will be assigned to the Githio cluster, whereas for the municipality of Krokeon the index $I$ will be calculated since it has common borders with Skala. Similarly, the municipality of Therapnon will be unionised to Skala, while for the remaining two units (Elous and Geronyron) adjacent to the second center, the index $I$ will determine which of the two clusters (Skalas or Niaton) they will join. Lakonia unionises with the municipalities of Zaraka and Molaon. Calculated values of the index $I$
indicate that Geronyron and Krokeon should join Skalas cluster, whereas Elous that of Niaton. As a result, all municipalities are assigned to the three clusters until, according to the termination condition of the Step 5, the join criterion is satisfied.

Figure 2. Initial cluster centers.

After the allocation of the remaining polygons (Step 6) the three clusters are initially defined and the up to this point (common border and distance of the centroids) electoral map of Lakonia is depicted in Figure 3. The synthesis as well as the total population of the three defined clusters is shown in Table 1.

A first and obvious conclusion deriving from the table is that the district population distribution needs to be optimised in order to achieve the desirable level of homogeneity between the newly designed clusters. Improvement of the clusters quantitative characteristics is accomplished by means of standard deviation of population per seat calibration according to Steps 7 and 8.

The standard deviation of the three clusters of Lakonia is 10.981. Potential reallocations can be considered for the polygons of Krokeon and Spartis to the cluster of Githio that is the only one meeting the defined electoral standard. In this respect, transferring the Krokeon municipality is initially examined since it shares the longest common border with the specific cluster. In this case, the standard deviation drops to 8.110, therefore the movement is made permanent and the population of the cluster rises to 23.808 people. When the Spartis polygon is transferred standard deviation increases to 8.212 and thus joining is not applied.
Since there are no possible movements left to the cluster of Githio, the cluster with the maximum population (Skalas) is examined whether one of its polygon units can be separated. Starting from peripheral polygons, in ascending order, those of Skalas, Geronyron and Spartis
are selected. Their staged examination, results to the reallocation of Skalas and Geronyron to the cluster of Githio, decreasing standard deviation to 844 people. The finally resulting area scheme according to our methodology is depicted in Figure 4 while the respective populations of the three redefined electoral districts are shown in Table 2.

![Figure 4. New electoral districts of Lakonia.](image)

**Table 2.** Population size of the new electoral districts of Lakonia.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1613</td>
<td>NIATON</td>
<td>2557</td>
<td>31860</td>
</tr>
<tr>
<td>1607</td>
<td>ZARAKA</td>
<td>1696</td>
<td></td>
</tr>
<tr>
<td>1610</td>
<td>MOLAON</td>
<td>5472</td>
<td></td>
</tr>
<tr>
<td>1606</td>
<td>ELOUS</td>
<td>5992</td>
<td></td>
</tr>
<tr>
<td>1611</td>
<td>MONEMBASIAS</td>
<td>3950</td>
<td></td>
</tr>
<tr>
<td>1602</td>
<td>ASOPOU</td>
<td>3666</td>
<td></td>
</tr>
<tr>
<td>1603</td>
<td>VOION</td>
<td>7802</td>
<td></td>
</tr>
<tr>
<td>1621</td>
<td>ELAFONISOU</td>
<td>725</td>
<td></td>
</tr>
<tr>
<td>1608</td>
<td>THERAPNON</td>
<td>2999</td>
<td>31075</td>
</tr>
<tr>
<td>1614</td>
<td>OINOUNTOS</td>
<td>2649</td>
<td></td>
</tr>
<tr>
<td>1619</td>
<td>SPARTIS</td>
<td>16322</td>
<td></td>
</tr>
<tr>
<td>1622</td>
<td>KARION</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>1616</td>
<td>PELLANAS</td>
<td>3863</td>
<td></td>
</tr>
<tr>
<td>1612</td>
<td>MYSTRA</td>
<td>4582</td>
<td></td>
</tr>
<tr>
<td>1605</td>
<td>GITHIOU</td>
<td>7542</td>
<td>32761</td>
</tr>
<tr>
<td>1618</td>
<td>SMYNOUS</td>
<td>1537</td>
<td></td>
</tr>
<tr>
<td>1615</td>
<td>OITILOU</td>
<td>4985</td>
<td></td>
</tr>
<tr>
<td>1601</td>
<td>ANATOLIKIS MANIS</td>
<td>2024</td>
<td></td>
</tr>
<tr>
<td>1609</td>
<td>KROKEON</td>
<td>2871</td>
<td></td>
</tr>
<tr>
<td>1617</td>
<td>SKALAS</td>
<td>6919</td>
<td></td>
</tr>
<tr>
<td>1604</td>
<td>GERONYRON</td>
<td>2034</td>
<td></td>
</tr>
<tr>
<td>1620</td>
<td>FARIDOS</td>
<td>4649</td>
<td></td>
</tr>
</tbody>
</table>
5. CONCLUDING REMARKS

The need for objective and unbiased organization of space, based on a set of predefined quantitative criteria, is intense. Spatial clustering of administrative regions is a critical planning issue with an extremely wide field of applications. At the same time is politically sensitive in the sense that it can be applied in favour or against one or more regions. The proposed methodology as it was realised by the SPIRAL algorithm, can improve the geographical compactness of a region through simple and comprehensive steps and towards the goal of homogeneity.

Both the conceptual spatial clustering framework and the overall regionalization approach allow the simultaneous consideration of variables that refer to and ensure the adjacency and homogeneity of constitutional units. Furthermore, SPIRAL algorithm provides the ability to optimize the outcome by analysing and smoothing the criterion-related deviation of the final clusters. In this respect, the question “how many regions” is effectively and alternatively confronted under specific constraints and requirements.

Considering the algorithm’s spatial nature, the exploitation of GIS during the execution of the algorithm, is essential not only due to the topological information processed, but also to their increased spatial analysis and cartographic representation functionality. The extended spatial databases needed remain a critical problem parameter that significantly contributes to some time-consuming issues of the approach.

Improvements with respect to the user interface and the integration of the SPIRAL algorithm’s steps to a unified open-source system platform should be considered as a fundamental prerequisite for its further utilisation. According to the literature, as GIS become more involved in spatial problem solving procedures, the overall process is significantly improved through the fluent analysis of a set of hypotheses and scenarios and the definition of solutions exhibiting advanced levels of effectiveness and efficiency.

With respect to the algorithm, an intervention that will improve the region-building outcome is the consideration of the second order contiguity in the determination of the initial clusters and the calculation of the index of Unification $I$. In this manner, more complicated spatial problems can be resolved since the analysis of space will be more thorough. In a similar framework, more than one criterion can be exploited and interrelated (service area size, average distance travelled and structural capacity to name a few) widening its spectrum of applications. As a result decreased processing times will be achieved as well as meaningful and useful regions will be obtained.

In conclusion, SPIRAL algorithm succeeded in defining homogenous spatial clusters according to a specific criterion with the resulting regions meeting the set constraint. The fictitious case study underlined the importance of adjacency in the determination of administrative regions while reassuring the rational formulation of spatial patterns. However, in using methods like the one presented, decision makers need to be aware of several practical issues that stem from and are reflected to the input data variable(s). Their careful and justified selection is not only essential for but a truly sine qua non condition.

REFERENCES


