A Multi-Objective approach to visualize proportions and similarities between individuals by rectangular maps

Emilio Carrizosa[1], Vanesa Guerrero[1], and Dolores Romero Morales[2]

[1]Instituto de Matemáticas de la Universidad de Sevilla (IMUS), Seville, Spain
{ecarrizosa, vguerrero}@us.es
[2]Copenhagen Business School, Frederiksberg, Denmark
drm.eco@cbs.dk

Abstract

In this paper we address the problem of visualizing the proportions and the similarities attached to a set of individuals. We represent this information using a rectangular map, i.e., a subdivision of a rectangle into rectangular portions so that each portion is associated with one individual, their areas reflect the proportions, and the closeness between portions represents the similarity between the individuals. By considering the most similar individuals as adjacent, we seek to represent adjacent individuals as adjacent portions in the rectangular map. Due to the impossibility of satisfying both area and adjacency requirements, this visualization problem is formulated as a three-objective Mixed Integer Nonlinear Problem. The first objective seeks to maximize the number of true adjacencies that the rectangular map is able to reproduce, the second one is to minimize the number of false adjacencies that the rectangular map adds, and the last one is to minimize the total deviation of the areas of the portions in the rectangular map from the given proportions. To guide the location of the rectangles, we have designed a tailored MultiDimensional Scaling for building rectangular maps. We study the tradeoff between the three objectives by solving the problem with their weighted summation. Our numerical results demonstrate that it is possible to provide a collection of rectangular maps with different tradeoffs between an accurate representation of the proportions by areas versus an accurate representation of the similarities by adjacencies.

Keywords: Nonlinear Programming, Mixed Integer Programming, Visualization, MultiDimensional Scaling

1 Introduction

In the Big Data era, it is critical to enable analysts to observe data using appropriate visualization tools, see [35, 47]. A natural and frequent task is to depict a set of individuals $V = \{v_1, \ldots, v_N\}$, to which there are attached some statistical values represented as proportions, $\omega = (\omega_1, \ldots, \omega_N)$, where without loss of generality we assume that $\sum_{r=1}^{N} \omega_r = 1$, see, e.g., [43]. Market share, vote intention or population rates, just to name a few, are usual examples. In order to visualize proportions, a common approach is to consider a bounded region of the plane and to subdivide it into portions $\mathbf{P} = (P_1, \ldots, P_N)$ of common shape whose areas represent the proportions. Well-known visualization tools for this kind of data are the classic pie or fan charts, and the treemap [30, 41], see Figure 1. In this kind of representations, holes are not allowed, thus, receiving the name of planar space-filling visualization maps.

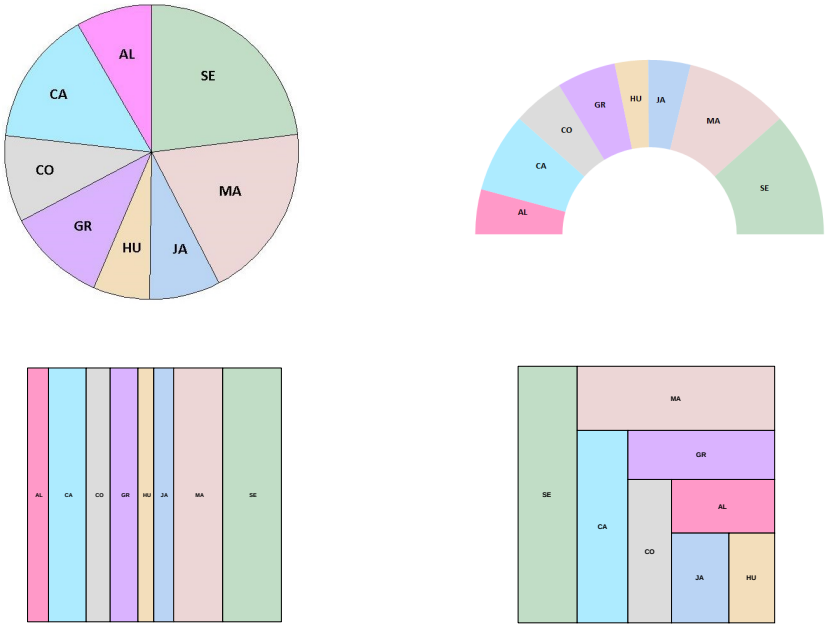


Figure 1: Examples of planar space-filling visualization maps: pie chart, fan chart, and two treemaps

A planar space-filling map to visualize the proportions attached to individuals in a bounded set Ω of the plane can be found by constructing the portions of the desired area and putting them together to fill Ω. This is straightforward in the case of the pie or fan charts: for a permutation $\sigma(1), \sigma(2), \ldots, \sigma(N)$ of the indices $1, 2, \ldots, N$, portions of areas $\omega_{\sigma(1)}, \omega_{\sigma(2)}, \ldots, \omega_{\sigma(N)}$ are placed sequentially in Ω. The only freedom in such planar space-filling visualization maps is thus the choice of the permutation. For the case of *rectangular maps*, [8], where Ω is the unit square and portions are rectangles, the same approach can be used, where the rectangular portions go all the way from North to South (or, by rotation, from West to East). This is illustrated in Figure 1, bottom-left visualization map. While pie and fan charts only admit different sequential arrangements, rectangular maps allow more freedom than the choice of a permutation. This is

the case for treemaps, which are rectangular maps depicting the hierarchy in weighted trees. This is illustrated in Figure 1, bottom-right visualization map.

The flexible layout offered by rectangular maps is also desirable when visualizing a similarity attached to the set of individuals. The nature of the similarity can be diverse. For instance, there exist proposals to improve treemap representations by taking into account the spatial closeness of data, [9, 11, 22, 49, 52]. But there are many other examples of similarities, which are not necessarily spatial, as we will see in our experimental section. A common approach in the literature has been to represent *close* individuals as *close* portions in the visualization map, see [19, 26] and references therein. A threshold in the similarities is established, and two individuals v_r and v_s are considered *adjacent*, $e_{rs} = 1$, if their similarity is higher than the threshold, otherwise $e_{rs} = 0$. In this paper, we build rectangular maps to visualize proportions $\omega = (\omega_1, \ldots, \omega_N)$ and similarities between the individuals, as measured by the adjacency matrix $E = (e_{rs})$. Hereafter, we denote $G = (V, E, \omega)$ and refer indistinctly to individuals with proportions and similarities attached or weighted graph throughout the paper. Similarly, we denote $G_{\mathbf{P}} = (V, E_{\mathbf{P}}, area_{\mathbf{P}})$, the weighted graph associated with the rectangular map, in such a way that $(v_r, v_s) \in E_{\mathbf{P}}$ if portions P_r and P_s are adjacent, i.e., their borders intersect in an infinite amount of points. The areas of the rectangles in the rectangular map correspond to the node weights in $G_{\mathbf{P}}$.

Visualizing statistical values (proportions) associated with the geographical regions (individuals) has been studied in cartography, [48]. Representing geographical regions with relatively simple shapes, such as rectangles, and whose areas represent, for instance, population rates or vote intention, as well as relative positions between regions are maintained, is important for many disciplines. This is, for instance, the case in Demography, Social Sciences, Biology, and even in Journalism. *Rectangular cartograms*, which were first introduced in [38], have been the most fruitful area in the development of rectangular maps, see for example [21, 27, 33, 48]. However, in general, one cannot guarantee the existence of a rectangular map that satisfies area and adjacency requirements on the rectangles, see [33], especially when the graph G to represent is not planar. See [5, 10] for further complexity results on rectangular maps. This impossibility has motivated a soft modeling of the criteria. However, the problem remains hard. The most common approach in cartography to reduce the number of possible layouts for rectangular maps is to consider that each rectangle must contain a point, which is the centroid of the geographical region, [19, 26, 27, 52]. However, this approach does not apply when the individuals are not geographical regions.

Although not on visualization, related problems can be found in other fields. The case in which there are no proportions (weights) attached to the individuals, and the graph G is planar, has been studied in the literature and it has many applications, for instance, in Very Large Scale Integration circuits design [6, 46]. The usual approach there is to find a *rectangular dual of a planar graph*, which consists of a subdivision of the unit square in such a way that each vertex (individual) corresponds to a different rectangle in the subdivision and, if v_r and v_s are linked, then the corresponding portions P_r and P_s are adjacent in the subdivision. Some characterizations of planar graphs that admit a rectangular dual can be found in [10, 15, 32]. Rectangular duals are also related with facility layout, [7, 18, 25, 29, 40], and graph drawing, [37, 41].

In this paper, we propose a multi-objective program to build rectangular maps which simultaneously optimizes the usually called "topology measure" or "adjacency measure" and the "area deviation" for weighted graphs G, not necessarily planar. These goodness measures as

well as others can be found in the literature, see [4, 27], although they have not been properly optimized to obtain a variety of maps. We consider the unit square Ω split into K rows and L columns, each cell representing thus a $100/(K \times L)\%$ of the total area of Ω. This grid structure has been also considered in other approaches to construct planar space-filling visualization maps, see [51], and it allows to easily measure areas, and it simplifies the notion of adjacency, since two portions are adjacent if they touch in, at least, one cell. The aspect ratios of the rectangles are also bounded by the minimum of the cells width or height, avoiding very narrow portions if the grid is coarse enough.

We formulate the problem of building rectangular maps as a multi-objective Mixed Integer Nonlinear Programming (MINLP). Such MINLP is a difficult problem, in part due to its non-linearities (we have the product of 0–1 variables to model the adjacencies between portions). However, it is easily linearized and transformed into a linear Mixed Integer Program (MIP). Nevertheless, such MIP is still be too demanding in terms of running times. We have designed a tailored MultiDimensional Scaling (MDS), [34], for rectangular maps. This MDS can handle *any* set of individuals with proportions and similarities attached, not only geographical. Similarly as in [36], our MDS builds a set of points, called *locating points*, taking into account that those points will belong to disjoint rectangles of certain area and also that distances between the points must resemble the similarity between the individuals.

The remainder of the paper is structured as follows. In Section 2 we introduce the multi-objective program to build rectangular maps. In Section 3 we formulate the problem as an MINLP, which is linearized and transformed into a linear MIP. In Section 4, we present the formulation for our tailored MDS for rectangular maps to obtain the locating points. Section 5 is the experimental section. Our approach is illustrated using three examples of diverse nature: the first one consists of the proportion of people in each blood group in the U.S. and the compatibility between the groups, while the other two are cartographic applications. An analysis of the trade-off between the objectives in our linear MIP, namely, the true adjacencies reproduced, the false adjacencies added, and the total area deviations, is provided. Section 6 concludes the paper with a summary and lines for future research.

2 The problem

An ideal rectangular map to visualize the proportions attached to individuals and adjacencies, $G = (V, E, \omega)$, would consist of a subdivision $\mathbf{P}$ of Ω into portions $P_1, \ldots, P_N$ of rectangular shape so that the following conditions hold

(C1) $\bigcup_{r=1}^{N} P_r = \Omega$

(C2) $int(P_r) \cap int(P_s) = \emptyset,\ \forall r, s = 1, \ldots, N, r \neq s$

(C3) $\text{area}(P_r) = \omega_r,\ \forall r = 1, \ldots, N$

(C4) $E = E_{\mathbf{P}}$

(C5) P_r is a rectangle, made up of cells of the (K, L)-grid in which $\Omega = [0, 1] \times [0, 1]$ is divided.

We propose the Multi-Objective Rectangular Map ($MORM$) formulation to obtain (K, L)-rectangular maps to visualize the set of individuals V, while pursuing an accurate representation of the proportions ω by areas and an accurate representation of the adjacencies E:

$$
\begin{array}{lll}
\max & |E \cap E_{\mathbf{P}}| & \\
\min & |E^c \cap E_{\mathbf{P}}| & \\
\min & \sum_{r=1}^{N} |\omega_r - area(P_r)| & (MORM) \\
\text{s.t.} & \mathbf{P} = (P_1, \ldots, P_N) \text{ satisfying (C1), (C2), (C5).} &
\end{array}
$$

The first two objectives model the closeness between E and $E_{\mathbf{P}}$, i.e. (C4), by means of the cardinality of the sets $E \cap E_{\mathbf{P}}$ and $E^c \cap E_{\mathbf{P}}$, where E^c denotes the complement of E. The third objective comes from considering the condition (C3), $area(P_r) = w_r$, as a soft constraint, and minimizing the sum of the deviations. These objectives are combined into one single objective following the standard additive scalarization [20]. By sampling the aggregation parameters space, different Pareto-optimal solutions are obtained. This way, different aggregations yield different rectangular maps, with a different trade-off between the criteria involved. As far as we are aware, this three-objective approach has not been studied before in the literature from an optimization point of view.

Figure 3 illustrates the concept of (K, L)-rectangular map and the three objectives under consideration, using as G the weighted graph plotted in Figure 2, where $N = 6$, $|E| = 9$ and $\omega = (0.3, 0.15, 0.1, 0.15, 0.1, 0.2)$. Figure 3a represents G as a $(5, 10)$-rectangular map, where the $K = 5$ rows are numbered from top to bottom and the $L = 10$ columns from left to right. We may observe that 8 out of the 9 true adjacencies, i.e., the adjacencies in E, are reproduced by $E_{\mathbf{P}}$, which are shown as black edges in the graph in Figure 3b. There is only one true adjacency missing in $E_{\mathbf{P}}$, represented as a dashed red edge in Figure 3b: v_3 and v_4 are adjacent in G but their associated rectangles P_3 and P_4 are not in the rectangular map. (Note that if two cells touch only in a corner, they are not considered adjacent.) Finally, the rectangular map adds a false adjacency, i.e., an adjacency which was not in E, drawn as a dashed red edge in Figure 3c: v_2 and v_4 are not adjacent in G but P_2 and P_4 are in the rectangular map. Finally, and with respect to the weights, the rectangular map approximates them. For instance, v_4 has associated a weight equal to $\omega_4 = 0.15$, while the area of P_4 is equal to $4/50 = 0.08$.

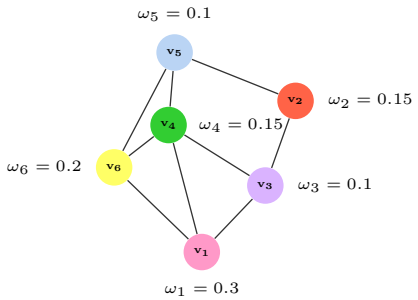
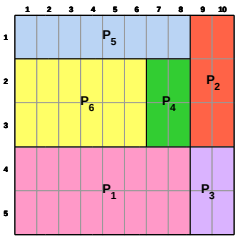
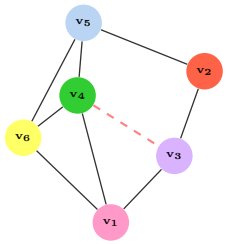


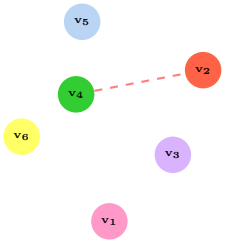
Figure 2: Example graph $G = (V, E)$; $N = 6$, $|E| = 9$, $\omega = (0.3, 0.15, 0.1, 0.15, 0.1, 0.2)$



(a) $(5, 10)$-rectangular map of G in Figure 2



(b) $E \cap E_{\mathbf{P}}$ in black. $E \setminus E_{\mathbf{P}}$ in dashed red.



(c) $E^c \cap E_{\mathbf{P}}$ in dashed red.

Figure 3: A $(5, 10)$-rectangular map for G; $|E \cap E_{\mathbf{P}}| = 8$, $|E^c \cap E_{\mathbf{P}}| = 1$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.24$.

3 An MINLP formulation for building rectangular maps

In this section we formulate Problem $(MORM)$ as a three-objective MINLP. The three objectives are aggregated in the usual way, by considering their weighted sum. We present the decision variables in Section 3.1, the objective function in Section 3.2, and the constraints in Section 3.3. The problem is then written as a linear MIP in Section 3.4. In what follows, indices r and s are used for portions, i for rows of the grid, j for columns, and l to indicate adjacency from side l, where $l = 1$ refers to adjacency from above, $l = 2$ below, $l = 3$ right, and $l = 4$ left. Finally, we use the initials N for North, S for South, E for East, and W for West.

3.1 Decision variables

We use 0–1 as well as continuous decision variables. We have two types of variables associated with (portion, cell) combinations, namely, the cell assignment variables x_{rij} and the corner ones y^{NW}_{rij} and y^{SE}_{rij}, where NW indicates the North West corner and SE the South East one. We have variables z^l_{rs}, associated with pair of portions, to measure adjacency from side l and which allow to count the adjacencies between pair of portions in the rectangular map. Thus,

$$
\begin{aligned}
x_{rij} &= \begin{cases} 1 & \text{if cell } (i,j) \text{ belongs to portion } P_r \\ 0 & \text{otherwise} \end{cases} \\
y^{NW}_{rij} &= \begin{cases} 1 & \text{if cell } (i,j) \text{ is the NW corner of portion } P_r \\ 0 & \text{otherwise} \end{cases} \\
y^{SE}_{rij} &= \begin{cases} 1 & \text{if cell } (i,j) \text{ is the SE corner of portion } P_r \\ 0 & \text{otherwise} \end{cases} \\
z^l_{rs} &= \begin{cases} 1 & \text{if portion } P_s \text{ is adjacent to portion } P_r \text{ from side } l \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Finally, we have nonnegative auxiliary variables φ_r and ψ_r, to linearize the area deviation $|\omega_r - area(P_r)|$, i.e., $|\omega_r - area(P_r)| = \varphi_r + \psi_r$ and $\omega_r - area(P_r) = \varphi_r - \psi_r$.

We illustrate these variables using the example in Figure 2. For instance, in the rectangular map in Figure 3a, portion P_4 has four cells defined by $x_{427} = x_{428} = x_{437} = x_{438} = 1$. The cell $(2, 7)$ is the NW corner of portion P_4 and cell $(3, 8)$ is its SE corner, thus $y^{NW}_{427} = y^{SE}_{438} = 1$. Moreover, portion P_4 has four adjacent portions: P_1 from below, P_2 to the right, P_5 from above and P_6 to the left. Thus, $z^2_{41} = z^3_{42} = z^1_{45} = z^4_{46} = 1$. The remaining binary variables of the form x_{4ij}, y^{NW}_{4ij}, y^{SE}_{4ij} and z^l_{4s} are zero. As noted in the introduction, the area of P_4 is equal to $4/50 = 0.08$ while $\omega_4 = 0.15$, therefore $\varphi_4 = 0.07$ and $\psi_4 = 0$.

3.2 Objective function

Because of the definition of the variables, it is straightforward to see that the three objectives in Problem $(MORM)$ (written in maximization form) are, respectively,

$$\begin{aligned} & \sum_{\substack{r,s=1\dots N \\ l=1,\dots,4 \\ (r,s)\in E}} z^l_{rs} \\ - & \sum_{\substack{r,s=1\dots N \\ l=1,\dots,4 \\ (r,s)\in E^c}} z^l_{rs} \\ - & \sum_{r=1,\dots,N} (\varphi_r + \psi_r)\,. \end{aligned}$$

The additive scalarization of the objectives is thus given by

$$\lambda_1 \sum_{\substack{r,s=1\dots N \\ l=1,\dots,4 \\ (r,s)\in E}} z^l_{rs} - \lambda_2 \sum_{\substack{r,s=1\dots N \\ l=1,\dots,4 \\ (r,s)\in E^c}} z^l_{rs} - \lambda_3 \sum_{r=1,\dots,N} (\varphi_r + \psi_r)\,, \tag{1}$$

to be maximized, for fixed scaling nonzero vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$, $\lambda_t \geq 0,\ t = 1, 2, 3$.

3.3 Constraints

We now write the constraints in Problem $(MORM)$ using the decision variables above, and give a brief explanation of each group of constraints. Note that condition (C5) is satisfied thanks to the definition of the x-variables and constraint (22); conditions (C1) and (C2) follow from constraint (23), which ensure that each cell is assigned to exactly one portion; condition (C4) is modeled through the first and second objectives plus constraints (2)-(15). Finally, the soft version of condition (C3) is modeled through constraint (25) jointly with the third objective. The remaining constraints model the correctness of the definition of variables and the symmetry of the problem.

$$\sum_{\substack{i=2,\dots,K \\ j=1,\dots,L}} x_{rij} \cdot x_{s\,i-1\,j} \geq z^1_{rs}, \qquad r, s = 1, \dots, N,\ r \neq s \tag{2}$$

$$\sum_{\substack{i=1,\dots,K-1 \\ j=1,\dots,L}} x_{rij} \cdot x_{s\,i+1\,j} \geq z^2_{rs}, \qquad r, s = 1, \dots, N,\ r \neq s \tag{3}$$

$$\sum_{\substack{i=1,\dots,K\\ j=1,\dots,L-1}} x_{rij}\cdot x_{s\,i\,j+1} \geq z^3_{rs}, \qquad r,s=1,\dots,N,\ r\neq s \tag{4}$$

$$\sum_{\substack{i=1,\dots,K\\ j=2,\dots,L}} x_{rij}\cdot x_{s\,i\,j-1} \geq z^4_{rs}, \qquad r,s=1,\dots,N,\ r\neq s \tag{5}$$

$$x_{rij}+x_{s\,i-1\,j}\leq z^1_{rs}+1 \qquad i=2,\dots,K,\ j=1,\dots,L,\ r,s=1,\dots,N,\ r\neq s \tag{6}$$

$$x_{rij}+x_{s\,i+1\,j}\leq z^2_{rs}+1 \qquad i=1,\dots,K-1,\ j=1,\dots,L,\ r,s=1,\dots,N,\ r\neq s \tag{7}$$

$$x_{rij}+x_{s\,i\,j+1}\leq z^3_{rs}+1 \qquad i=1,\dots,K,\ j=1,\dots,L-1,\ r,s=1,\dots,N,\ r\neq s \tag{8}$$

$$x_{rij}+x_{s\,i\,j-1}\leq z^4_{rs}+1 \qquad i=1,\dots,K,\ j=2,\dots,L,\ r,s=1,\dots,N,\ r\neq s \tag{9}$$

$$z^1_{rs}\leq 1-x_{r1j} \qquad j=1,\dots,L,\ r,s=1,\dots,N,\ r\neq s \tag{10}$$

$$z^2_{rs}\leq 1-x_{rKj} \qquad j=1,\dots,L,\ r,s=1,\dots,N,\ r\neq s \tag{11}$$

$$z^3_{rs}\leq 1-x_{riL} \qquad i=1,\dots,K,\ r,s=1,\dots,N,\ r\neq s \tag{12}$$

$$z^4_{rs}\leq 1-x_{ri1} \qquad i=1,\dots,K,\ r,s=1,\dots,N,\ r\neq s \tag{13}$$

$$z^1_{rs}=z^2_{sr} \qquad r,s=1,\dots,N,\ r\neq s \tag{14}$$

$$z^3_{rs}=z^4_{sr} \qquad r,s=1,\dots,N,\ r\neq s \tag{15}$$

$$\sum_{\substack{i'\leq i\\ j'\leq j}} y^{NW}_{ri'j'}\geq x_{rij} \qquad i=1,\dots,K,\ j=1,\dots,L,\ r=1,\dots,N \tag{16}$$

$$\sum_{\substack{i'\geq i\\ j'\geq j}} y^{SE}_{ri'j'}\geq x_{rij} \qquad i=1,\dots,K,\ j=1,\dots,L,\ r=1,\dots,N \tag{17}$$

$$\sum_{\substack{i=1,\dots,K\\ j=1,\dots,L}} y^{NW}_{rij}=1 \qquad r=1,\dots,N \tag{18}$$

$$\sum_{\substack{i=1,\dots,K\\ j=1,\dots,L}} y^{SE}_{rij}=1 \qquad r=1,\dots,N \tag{19}$$

$$y^{NW}_{rij}\leq x_{rij} \qquad i=1,\dots,K,\ j=1,\dots,L,\ r=1,\dots,N \tag{20}$$

$$y^{SE}_{rij}\leq x_{rij} \qquad i=1,\dots,K,\ j=1,\dots,L,\ r=1,\dots,N \tag{21}$$

$$x_{r\,i''\,j''}+1\geq y^{NW}_{rij}+y^{SE}_{r\,i'\,j'} \qquad i,i'=1,\dots,K,\ j,j'=1,\dots,L,\ r=1,\dots,N,\ i\leq i''\leq i',\ j\leq j''\leq j', \tag{22}$$

$$\sum_{r=1,\dots,N} x_{rij}=1 \qquad i=1,\dots,K,\ j=1,\dots,L \tag{23}$$

$$\sum_{\substack{i=1,\dots,K\\ j=1,\dots,L}} x_{rij}\geq 1 \qquad r=1,\dots,N \tag{24}$$

$$\sum_{\substack{i=1,\dots,K\\ j=1,\dots,L}} x_{rij}=KL\left(\omega_r-(\varphi_r-\psi_r)\right) \qquad r=1,\dots,N \tag{25}$$

$$x_{rij},\ y^{NW}_{rij},\ y^{SE}_{rij},\ z^{l}_{rs} \in \{0,1\} \qquad i=1,\dots,K,\ j=1,\dots,L, \qquad (26)$$
$$l=1,\dots,4,\ r,s=1,\dots,N,\ r\neq s$$
$$\varphi_r, \psi_r \geq 0 \qquad r=1,\dots,N. \qquad (27)$$

- Constraints (2)-(5) model the correctness of $z^l_{rs}=1$, i.e., if variable z^l_{rs} takes the value 1, then there must be two adjacent cells belonging to portions P_r and P_s respectively. For instance, if $z^1_{rs}=1$ then some (i,j) exists such that $x_{rij}=1$ and $x_{s\,i-1\,j}=1$, or equivalently constraint (2).
- Constraints (6)-(9) model the correctness of $z^l_{rs}=0$. For instance, if $z^1_{rs}=0$ and $x_{rij}=1$ then $x_{s\,i-1\,j}=0$.
- Constraints (10) and (11) say that there is no portion adjacent from above to those which are in row 1 and from below to those which are in row K, respectively.
- Constraints (12) and (13) say that there is no portion adjacent to the right to those which are in column L and to the left to those which are in column 1, respectively.
- Constraints (14) and (15) impose that some adjacency variables z are equal due to symmetry. For instance, constraint (14) imposes that if portion P_r is adjacent to P_s from above, then portion P_s is adjacent to portion P_r from below.
- Constraint (16) ensures that the NW corner of a portion P_r must be at the NW, that is, the NW cell is the cell in P_r with the lowest row and column indices, while (17) imposes that the SE corner of a portion P_r must be at the SE, that is, the SE cell is the cell in P_r with highest row and column indices. Moreover, (18) and (19) guarantee that each portion has only one NW and SE corner, respectively. Finally, (20) and (21) assure that both corners belong to the portion.
- Constraint (22) ensures that portions are rectangles, and thus condition (C4) is satisfied.
- Constraint (23) makes sure that each cell belongs to exactly one portion, i.e., (C1) and (C2) are satisfied.
- Constraint (24) avoids empty portions.
- Constraint (25) linearizes the absolute value $|\omega_r - area(P_r)|$.
- Finally, the type of the variables is modeled with constraints (26) and (27).

Note that two portions can be only adjacent on one side since they are rectangles. This condition follows from constraints (2)-(5) and (22).

3.4 Writing the problem as a linear MIP

Thus, given a weighted graph $G=(V,E,\omega)$, Pareto-optimal solutions to Problem $(MORM)$ are obtained by solving the following MINLP for different values of the scaling vector $\boldsymbol{\lambda}$:

$$\begin{aligned} &\max && (1)\\ &\text{s.t.} && (2)-(27). \end{aligned}$$

Observe that nonlinearities in this formulation appear only in products between 0–1 variables, namely in constraints (2)-(5). Therefore, by introducing auxiliary variables and additional constraints these products can be eliminated and the resulting problem is then a linear MIP. We proceed in the usual way by defining variables u as follows:

$$\begin{aligned} u^1_{rsij} &= x_{rij} \cdot x_{s\,i-1\,j} \\ u^2_{rsij} &= x_{rij} \cdot x_{s\,i+1\,j} \\ u^3_{rsij} &= x_{rij} \cdot x_{s\,i\,j+1} \\ u^4_{rsij} &= x_{rij} \cdot x_{s\,i\,j-1}. \end{aligned}$$

Note that u^l_{rsij} indicates whether portions r and s are adjacent at cell (i,j) at side l. Thus, constraints (2)-(5) become

$$\sum_{\substack{i=2,\ldots,K \\ j=1,\ldots,L}} u^1_{rsij} \geq z^1_{rs}, \qquad r,s=1,\ldots,N,\ r\neq s \tag{28}$$

$$\sum_{\substack{i=1,\ldots,K-1 \\ j=1,\ldots,L}} u^2_{rsij} \geq z^2_{rs}, \qquad r,s=1,\ldots,N,\ r\neq s \tag{29}$$

$$\sum_{\substack{i=1,\ldots,K \\ j=1,\ldots,L-1}} u^3_{rsij} \geq z^3_{rs}, \qquad r,s=1,\ldots,N,\ r\neq s \tag{30}$$

$$\sum_{\substack{i=1,\ldots,K \\ j=2,\ldots,L}} u^4_{rsij} \geq z^4_{rs}, \qquad r,s=1,\ldots,N,\ r\neq s. \tag{31}$$

Finally, to model variables u^l correctly we need to link them to variables x by adding the following set of constraints:

$$u^1_{rsij} \leq x_{rij} \qquad i=1,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{32}$$

$$u^1_{rsij} \leq x_{s\,i-1\,j} \qquad i=2,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N\,,r\neq s \tag{33}$$

$$x_{rij} + x_{s\,i-1\,j} \leq u^1_{rsij} + 1 \qquad i=2,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{34}$$

$$u^2_{rsij} \leq x_{rij} \qquad i=1,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{35}$$

$$u^2_{rsij} \leq x_{s\,i+1\,j} \qquad i=1,\ldots,K-1,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{36}$$

$$x_{rij} + x_{s\,i+1\,j} \leq u^2_{rsij} + 1 \qquad i=1,\ldots,K-1,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{37}$$

$$u^3_{rsij} \leq x_{rij} \qquad i=1,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{38}$$

$$u^3_{rsij} \leq x_{s\,i\,j+1} \qquad i=1,\ldots,K,\ j=1,\ldots,L-1,\ r,s=1,\ldots,N,\ r\neq s \tag{39}$$

$$x_{rij} + x_{s\,i\,j+1} \leq u^3_{rsij} + 1 \qquad i=1,\ldots,K,\ j=1,\ldots,L-1,\ r,s=1,\ldots,N,\ r\neq s \tag{40}$$

$$u^4_{rsij} \leq x_{rij} \qquad i=1,\ldots,K,\ j=1,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{41}$$

$$u^4_{rsij} \leq x_{s\,i\,j-1} \qquad i=1,\ldots,K,\ j=2,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{42}$$

$$x_{rij} + x_{s\,i\,j-1} \leq u^4_{rsij} + 1 \qquad i=1,\ldots,K,\ j=2,\ldots,L,\ r,s=1,\ldots,N,\ r\neq s \tag{43}$$

$$u^l_{rsij} \in \{0,1\} \qquad i=1,\ldots,K,\ j=1,\ldots,L,\ l=1,\ldots,4,\ r,s=1,\ldots,N,\ r\neq s. \tag{44}$$

Hence, Pareto-optimal solutions to $(MORM)$ can be generated by solving the linear MIP

$$\begin{array}{ll} \max & (1) \\ \text{s.t.} & (6)-(44) \end{array} \qquad (RM_{\boldsymbol{\lambda}})$$

for different values of the scaling vector $\boldsymbol{\lambda}$. However, even the very small instances we tested in our preliminary tests turned out to be too hard for standard MIP optimizers.

Borrowing the idea from other visualization procedures, such as the Recmap, [27], we guide the location of the different portions P_r by forcing them to contain so-called locating points $q_r \in [0,1] \times [0,1]$, $r = 1, \ldots, N$, built by an external procedure, as the one described in Section 4.

Hence, assuming a vector $\mathbf{q} = (q_1, \ldots, q_N)$ of locating points is given, the resolution of $(RM_{\boldsymbol{\lambda}})$ is replaced by solving the same problem with locating points

$$\begin{array}{ll} \max & (1) \\ \text{s.t.} & (6)-(44) \\ & q_r \in P_r, \ r = 1, \ldots, N. \end{array} \qquad (RM_{\boldsymbol{\lambda},\mathbf{q}})$$

Thanks to these extra constraints, some assignment variables x can be fixed, and the problem becomes easier, as our numerical results will show. However, we stress that the constraints related to the locating points are heuristic, and thus we cannot guarantee that the optimal solution obtained is also optimal to $(RM_{\boldsymbol{\lambda}})$, and thus Pareto-optimal to $(MORM)$. In the following section we describe an approach to judiciously build locating points.

4 MultiDimensional Scaling for rectangular maps

In this section we propose a tailored MultiDimensional Scaling (MDS), [34], for rectangular maps to find the locating points. Our MDS takes into account that the locating points are to be used by $(RM_{\boldsymbol{\lambda},\mathbf{q}})$, i.e., they have to lie in the unit square Ω and be part of the non-overlapping rectangular portions P_r whose areas are close to ω_r. See [31, 36] and references therein for other uses of MDS for planar visualization maps.

Let $D = (d_{rs})$ be the shortest path distance matrix between all nodes of graph $G = (V, E, \omega)$. We want to find N points $q_r = (q_r^1, q_r^2)$, contained in N rectangles defined by their NW and SE corners, (a_r^{NW}, b_r^{NW}) and (a_r^{SE}, b_r^{SE}) respectively. These rectangles, called in what follows MDS rectangles, are surrogate of the rectangular portions P_r in $(RM_{\boldsymbol{\lambda},\mathbf{q}})$, with some important differences. First, we do not impose that they are made of cells of the region Ω, avoiding the difficulties of the combinatorial part of the Problem $(RM_{\boldsymbol{\lambda}})$. Second, the MDS rectangles do not necessarily cover Ω. Third, they may overlap.

The locating points q_r are expected to be somehow central points of portions P_r in $(RM_{\boldsymbol{\lambda},\mathbf{q}})$, and thus the distance $\|q_r - q_s\|_1$ between locating points q_r and q_s should follow the same patterns than the distance d_{rs}. Hence, we impose that

$$\|q_r - q_s\|_1 \approx C d_{rs}, \qquad (45)$$

for some C, to be optimized. Observe that, contrary to standard MDS, in which Euclidean distances are used, here we propose to measure distances according to the ℓ_1 norm to better reflect the fact that we are dealing with rectangles with sides parallel to the coordinate axes.

We require two conditions to the MDS rectangles. We want the area of MDS rectangle r to approximate ω_r, i.e.,

$$(a_r^{SE} - a_r^{NW})(b_r^{NW} - b_r^{SE}) \approx \omega_r. \tag{46}$$

Moreover, we want the MDS rectangles not to overlap, but this is imposed as a soft constraint, forcing the area of each intersection being close to zero:

$$\max\left\{0, \min\{a_r^{SE}, a_s^{SE}\} - \max\{a_r^{NW}, a_s^{NW}\}\right\} \cdot \max\left\{0, \min\{b_r^{NW}, b_s^{NW}\} - \max\{b_r^{SE}, b_s^{SE}\}\right\} \approx 0. \tag{47}$$

With this notation, the MDS for rectangular maps is stated as the problem of finding rectangles, identified by their corner coordinates (a_r^{NW}, b_r^{NW}) and (a_r^{SE}, b_r^{SE}), and points q_r within minimal violation of soft constraints (45)-(47). This is expressed as the following nonlinear nonsmooth continuous optimization problem:

$$\begin{aligned}
\min\ & \gamma_1 \sum_{r,s=1}^{N} (d_{rs} - C\|q_r - q_s\|_1)^2 + \gamma_2 \sum_{i=r}^{N} \left(\left(a_r^{SE} - a_r^{NW}\right)\left(b_r^{NW} - b_r^{SE}\right) - \omega_r\right)^2 \qquad (48)\\
& + \gamma_3 \max\left\{0, \min\{a_r^{SE}, a_s^{SE}\} - \max\{a_r^{NW}, a_s^{NW}\}\right\} \cdot \max\left\{0, \min\{b_r^{NW}, b_s^{NW}\} - \max\{b_r^{SE}, b_s^{SE}\}\right\}\\
\text{s.t.}\\
& 0 \le a_r^{NW} \le q_r^1 \le a_r^{SE} \le 1, \qquad r = 1, \ldots, N \qquad (49)\\
& 0 \le b_r^{SE} \le q_r^2 \le b_r^{NW} \le 1, \qquad r = 1, \ldots, N \qquad (50)\\
& C > 0, \qquad (51)
\end{aligned}$$

where C is a scaling variable and γ_1, γ_2, $\gamma_3 \ge 0$ are scaling constants. Note that we can use a hyperbolic smoothing to approximate the absolute value and max and min functions

$$|t| \approx \sqrt{t^2 + \varepsilon},$$

$$\max\{t, t'\} = \frac{t + t' + |t' - t|}{2} \approx \frac{t + t' + \sqrt{(t' - t)^2 + \varepsilon}}{2},$$

$$\min\{t, t'\} = \frac{t + t' - |t' - t|}{2} \approx \frac{t + t' - \sqrt{(t' - t)^2 + \varepsilon}}{2},$$

where $\varepsilon > 0$.

Once $\mathbf{q} = (q_1, \ldots, q_N)$ is obtained, Problem $(RM_{\boldsymbol{\lambda}})$ is replaced by $(RM_{\boldsymbol{\lambda},\mathbf{q}})$. In case there exist $r \neq s$, where q_r and q_s belong to the same cell in the $K \times M$ cell grid, then $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ is infeasible. If this happens, several strategies are possible to recover a feasible problem. For instance, one could randomly perturb the locating points q_r and q_s until they lie in different cells. Alternatively, one can replace the constraint in $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ related with locating points by a weaker constraint of the form

$$q_r \in P_r, \ \forall r \in R, \tag{52}$$

where the set $R \subset \{1, \ldots, N\}$ is such that the different locating points belong to different cells. The latter approach is the one used in the experiments presented in Section 5.

5 Experimental results

In this section we illustrate our approach using three examples and a 16×16 grid. The examples are of diverse nature. The first one consists of visualizing the proportion of people in each blood group in the U.S. and the compatibility between the groups. The other two examples are cartographic applications. These two are planar graphs, while this is not the case for the first example. By sampling the scaling vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ different rectangular maps are obtained, with different trade-offs between the three objectives of $(MORM)$, namely, true adjacencies reproduced ($|E \cap E_{\mathbf{P}}|$), the false adjacencies added ($|E^c \cap E_{\mathbf{P}}|$) and the total area deviation ($\sum_{r=1}^{N} |\omega_r - area(P_r)|$). We have chosen $\lambda_1 \in \left\{0, \frac{1}{|E|}, \frac{5}{|E|}, \frac{10}{|E|}\right\}$, $\lambda_2 \in \left\{0, \frac{1}{|E^c|}, \frac{5}{|E^c|}, \frac{10}{|E^c|}\right\}$ and $\lambda_3 \in \{0, 1, 5, 10\}$, eliminating redundant ones due to homogeneity. The locating points are obtained by solving the MDS for rectangular maps given by (48)-(51) with $\gamma_1 = \gamma_3 = 1$ and $\gamma_2 = 1000$. Since it is a multimodal problem, a multistart with 50 runs is executed. All the MIPs involved have been written in AMPL, [24], and solved with CPLEX v12.6, [13], with a time limit of one hour, on a PC Intel® Core™ i7-2600K, 16GB of RAM. All the continuous nonlinear problems involved have been solved with the IPOPT solver, [50], through NEOS Server, [14].

The first example, `Blood`, consists of representing as a rectangular map the proportion of people in the U.S. in each blood group [44], taking into account the blood compatibility between donor and recipient. This example generates a weighted graph, the `Blood` graph, where nodes represent blood groups, and two groups v_r and v_s are considered similar (and thus adjacent) if group v_r can donate blood to group v_s. In the second example, `Netherlands`, the individuals are the provinces of The Netherlands, and the data represented are their (normalized) population, see [45]. The similarity measure considered is the geographical location, namely, two nodes are adjacent if the corresponding provinces are adjacent in the geographical map. The third example, `Germany`, is analogous to `Netherlands` but with a larger amount of individuals and adjacencies. We use Germany's geographical map, the data represented are their (normalized) geographical area, see [17], and again the geographical location as the similarity measure.

We follow an iterative procedure to generate the collection of rectangular maps. In each iteration, we solve $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ with a new vector of scaling parameters $\boldsymbol{\lambda}$. This $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ receives a solution pool which provides initial guesses to the current iteration. The solution pool consists of all the (16,16)-rectangular maps generated in previous iterations. In addition, we add to the pool the corresponding (8,8)-rectangular map. Solving $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ with $K = L = 8$ is a less demanding task, since both the number of decision variables and the number of constraints is much smaller than with $K = L = 16$. Indeed, for each example and each scaling vector, the optimal (8,8)-rectangular map was obtained within the one hour time limit. However, an 8×8 grid is a coarse grid and thus, we can find more than one locating point in the very same cell. This is the case in the `Netherlands` example, where q_4 and q_8 belong to cell (3,2) of the 8×8 grid. As we stated in Section 4, we build the set R where we have removed one of those points, q_8. Thus, $R = \{1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12\}$ and then Problem $(RM_{\boldsymbol{\lambda},\mathbf{q}})$, where constraint $q_r \in P_r$, $r = 1, \ldots, N$, is substituted by constraint (52), is solved firstly for an 8×8 grid and then for a 16×16 one.

Observe that we are making two approximations in our approach. First, we have imposed a time limit of one hour, so Problem $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ is not necessarily solved to optimality. Moreover, the choice of locating points as described in Section 4, although sensible, is heuristic. Hence, we cannot guarantee that the solutions provided by the optimizer when solving $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ are indeed

Pareto-optimal for $(MORM)$. However, they do give an approximation to the Pareto frontier by showing a trade-off between the different criteria involved in $(MORM)$, as the Figures 4-23 show.

The output of our experimental results is presented in Table 1 and Figures 4-23. In Table 1, the first five columns show the example features, i.e., number of individuals (nodes), number of adjacencies between the individuals (edges in the graph) and in its complement, and, to give an idea of the magnitude of the proportions, the maximum and minimum values (normalized), respectively. The next three columns, namely columns six, seven and eight, report for each rectangular map the value of the three objectives: true adjacencies reproduced, false adjacencies added and total area deviation. Finally, last column heads for Figures 4-23, representing the rectangular maps. Each figure contains three subfigures which represent, from left to right: the $(16, 16)$-rectangular map with the locating points; the graph G, where the true adjacencies reproduced $(E \cap E_{\mathbf{P}})$ have been highlighted in black and the rest of true adjacencies $(E \setminus E_{\mathbf{P}})$ in dashed red; the graph G, where the false adjacencies added by the rectangular map $(E^c \cap E_{\mathbf{P}})$ have been highlighted in dashed red.

For the `Blood` graph, we have three rectangular maps, see Figures 4-6. Among them, the maximum number of true adjacencies reproduced is 15 out of 19, while the minimum of false adjacencies added is 0, and the minimum total area deviation is 0.152. For the `Netherlands` graph, we have seven rectangular maps, see Figures 7-13. Among them, the maximum number of true adjacencies reproduced is 22 out of 22, while the minimum of false adjacencies added is 1, and the minimum total area deviation is 0.239. For the `Germany` graph, we have ten rectangular maps, see Figures 14-23. Among them, the maximum number of true adjacencies reproduced is 28 out of 28, while the minimum of false adjacencies added is 3, and the minimum total area deviation is 0.151.

6 Conclusions and future research

In this paper we have developed a new Multi-Objective approach to address the problem of representing a set of individuals, to which there are attached proportions and similarities, by means of rectangular maps. This kind of data can be modelled as a weighted graph and thus, our aim is to obtain rectangular maps where we maximize the adjacencies in the graph which are reproduced by the visualization map, while we minimize the false adjacencies added by the map as well as the deviation between individuals proportions and the area of the rectangles. Proceeding in the usual way, we consider the weighted summation of the objectives and formulate the problem as an MINLP, which is subsequently reformulated as a linear MIP. Due to its hard combinatorial structure, a tailored MultiDimensional Scaling has been designed to determine the relative positions of the rectangles in the map, and thus to reduce the number of possible layouts. Our approach has been illustrated using three examples, which differ in the planarity of the graph, the nature of the similarity and the number of individuals. Our results are competitive, and most of the true adjacencies (the ones in the original weighted graph) can be reproduced by the rectangular map introducing only a few false ones, and with low area deviations.

There are several interesting lines for future research. We are currently working on reinforcing the formulation of the Problems $(RM_{\boldsymbol{\lambda}})$ and $(RM_{\boldsymbol{\lambda},\mathbf{q}})$ with Reformulation-Linearization Techniques, [1, 2, 3, 39]. These techniques tighten the linear relaxation of MIPs, and, thus, are appealing to address finer grids, but also to develop new heurisitc approaches for building

rectangular maps. We are also working on other visualization tools. First, we are studying the problem of representing each node of the graph G by a connected union of grid cells, not necessarily with a rectangular shape [15, 16]. Having less rigid shapes than rectangles has two advantages, namely, the similarities between individuals can be represented more accurately, while better results in terms of area deviations can be achieved. Second, we would like to customize the technique of representing a set of individuals with attached proportions and similarities as a rectangular map to detect communities in graphs, [23], by analyzing the adjacencies represented in the rectangular map. Third, our method can also be applied to visualize hierarchical data, in which inside every rectangle a new rectangular map has to be represented by taking into account adjacencies with neighboring rectangles and its inner rectangular maps, [12, 28, 42].

Acknowledgement. *This research is funded in part by Projects MTM2012-36163 (Spain), P11-FQM-7603 and FQM-329 (Andalucía), all with EU ERD Funds.*

References

[1] W. P. Adams and H. D. Sherali. A tight linearization and an algorithm for zero-one quadratic programming problems. *Management Science*, 32(10):1274–1290, 1986.

[2] W. P. Adams and H. D. Sherali. Linearization strategies for a class of zero-one mixed integer programming problems. *Operations Research*, 38(2):217–226, 1990.

[3] W. P. Adams and H. D. Sherali. Mixed-integer bilinear programming problems. *Mathematical Programming*, 59(1-3):279–305, 1993.

[4] M. J. Alam, S. Kobourov, M. Scheneider, and S. Veeramoni. An experimental study of algorithms for cartogram generation. *Submitted to Algorithm Engineering and Experimentation (ALENEX'14)*, 2014.

[5] M.J. Alam, T. Biedl, S. Felsner, M. Kaufmann, S.G. Kobourov, and T. Ueckerdt. Computing cartograms with optimal complexity. *Discrete & Computational Geometry*, 50(3):784–810, 2013.

[6] M.F. Anjos and F. Liers. Global approaches for facility layout and VLSI floorplanning. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 849–877. Springer, 2012.

[7] M.F. Anjos and A. Vannelli. A new mathematical-programming framework for facility-layout design. *INFORMS Journal on Computing*, 18(1):111–118, 2006.

[8] T. Baudel and B. Broeksema. Capturing the design space of sequential space-filling layouts. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2593–2602, 2012.

[9] B. B. Bederson, B. Shneiderman, and M. Wattenberg. Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies. *ACM Transactions on Graphics*, 21(4):833–854, 2002.

[10] T. C. Biedl and B. Genç. Complexity of octagonal and rectangular cartograms. In *17th Canadian Conference on Computational Geometry*, pages 117–120, 2005.

[11] K. Buchin, D. Eppstein, M. Löffler, M. Nöllenburg, and R. I. Silveira. Adjacency-preserving spatial treemaps. In *Algorithms and Data Structures*, pages 159–170. Springer, 2011.

[12] S. Clémençon, H. De Arazoza, F. Rossi, and V.C. Tran. Hierarchical clustering for graph visualization. In *19th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2011)*, pages 227–232, 2011.

[13] IBM ILOG CPLEX. http://www.ilog.com/products/cplex/, 2014.

[14] J. Czyzyk, M.P. Mesnier, and J.J. More. The neos server. *IEEE Computational Science Engineering*, 5(3):68–75, 1998.

[15] M. de Berg, E. Mumford, and B. Speckmann. On rectilinear duals for vertex-weighted plane graphs. *Discrete Mathematics*, 309(7):1794–1812, 2009.

[16] M. de Berg, E. Mumford, and B. Speckmann. Optimal BSPs and rectilinear cartograms. *International Journal of Computational Geometry & Applications*, 20(02):203–222, 2010.

[17] Destatis, Statistisches Bundesamt. Area and population. www.destatis.de, 2015. Retrieved on: 2015-01-14.

[18] A. Drira, H. Pierreval, and S. Hajri-Gabouj. Facility layout problems: A survey. *Annual Reviews in Control*, 31(2):255–267, 2007.

[19] F. S. Duarte, F. Sikansi, F. M. Fatore, S. G. Fadel, and F. V. Paulovich. Nmap: A novel neighborhood preservation space-filling algorithm. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2063–2071, 2014.

[20] M. Ehrgott. A discussion of scalarization techniques for multiple objective integer programming. *Annals of Operations Research*, 147(1):343–360, 2006.

[21] D. Eppstein, E. Mumford, B. Speckmann, and K. Verbeek. Area-universal rectangular layouts. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, pages 267–276. ACM, 2009.

[22] D. Eppstein, M. van Kreveld, B. Speckmann, and F. Staals. Improved grid map layout by point set matching. In *IEEE Pacific Visualization Symposium (PacificVis)*, pages 25–32, 2013.

[23] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.

[24] R. Fourer, D. M. Gay, and B. W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Thomson/Brooks/Cole, 2003.

[25] R.L. Francis, J.A. White, and L.F. McGinnis. *Facility layout and location: an analytical approach*, volume 31. Prentice-Hall Englewood Cliffs, NJ, 1974.

[26] M. Ghoniem, M. Cornil, B. Broeksema, M. Stefas, and B. Otjacques. Weighted maps: treemap visualization of geolocated quantitative data. In *IS&T/SPIE Electronic Imaging*, pages 93970G–93970G. International Society for Optics and Photonics, 2015.

[27] R. Heilmann, D. A. Keim, C. Panse, and M. Sips. Recmap: Rectangular map approximations. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 33–40. IEEE Computer Society, 2004.

[28] I. Herman, G. Melançon, and M.S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, 2000.

[29] I. Jankovits, C. Luo, M. F. Anjos, and A. Vannelli. A convex optimisation framework for the unequal-areas facility layout problem. *European Journal of Operational Research*, 214(2):199–215, 2011.

[30] B. Johnson and B. Shneiderman. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *IEEE Conference on Visualization*, pages 284–291. IEEE, 1991.

[31] M. Klimenta and U. Brandes. Graph drawing by classical multidimensional scaling: new perspectives. In *Graph Drawing*, volume 7704, pages 55–66. Springer, 2013.

[32] K. Koźmiński and E. Kinnen. Rectangular duals of planar graphs. *Networks*, 15(2):145–157, 1985.

[33] M. van Kreveld and B. Speckmann. On rectangular cartograms. *Computational Geometry*, 37(3):175–187, 2007.

[34] J. B. Kruskal and M. Wish. *Multidimensional scaling*, volume 11. Sage, 1978.

[35] S. Liu, W. Cui, Y. Wu, and M. Liu. A survey on information visualization: recent advances and challenges. *The Visual Computer*, 30(12):1373–1393, 2014.

[36] X. Liu, Y. Hu, S. North, and H. Shen. Compactmap: A mental map preserving visual interface for streaming text data. In *IEEE International Conference on Big Data*, pages 48–55. IEEE, 2013.

[37] J. Owen-Smith, M. Riccaboni, F. Pammolli, and W.W. Powell. A comparison of U.S. and European university-industry relations in the life sciences. *Management Science*, 48(1):24–43, 2002.

[38] E. Raisz. The rectangular statistical cartogram. *Geographical Review*, pages 292–296, 1934.

[39] H. D. Sherali and L. Liberti. Reformulation-linearization technique for global optimization reformulation-linearization technique for global optimization. In *Encyclopedia of Optimization*, pages 3263–3268. Springer, 2009.

[40] H.D. Sherali, B.M.P. Fraticelli, and R.D. Meller. Enhanced model formulations for optimal facility layout. *Operations Research*, 51(4):629–644, 2003.

[41] B. Shneiderman. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on Graphics*, 11(1):92–99, 1992.

[42] B. Shneiderman and C. Dunne. Interactive network exploration to derive insights: Filtering, clustering, grouping, and simplification. In *Graph Drawing*, volume 7704, pages 2–18. Springer, 2013.

[43] I. Spence and S. Lewandowsky. Displaying proportions and percentages. *Applied Cognitive Psychology*, 5(1):61–77, 1991.

[44] Stanford Blood Center. Blood type in U.S. http://bloodcenter.stanford.edu/, 2014. Retrieved on: 2014-11-19.

[45] Statistics Netherlands. Population; gender, age, marital status and region, January 1. www.cbs.nl, 2013. Retrieved on: 2013-10-31.

[46] K. Tani, S. Tsukiyama, S. Shinoda, and I. Shirakawa. On area-efficient drawings of rectangular duals for vlsi floor-plan. *Mathematical Programming*, 52(1-3):29–43, 1991.

[47] J. Thomas and P.C. Wong. Visual analytics. *IEEE Computer Graphics and Applications*, 24(5):20–21, 2004.

[48] W. Tobler. Thirty five years of computer cartograms. *Annals of the Association of American Geographers*, 94(1):58–73, 2004.

[49] Y. Tu and H. Shen. Visualizing changes of hierarchical data using treemaps. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1286–1293, 2007.

[50] A. Wächter and L.T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

[51] M. Wattenberg. A note on space-filling visualizations and space-filling curves. In *IEEE Symposium on Information Visualization*, pages 181–186, 2005.

[52] J. Wood and J. Dykes. Spatially ordered treemaps. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1348–1355, 2008.

	GRAPH					RECTANGULAR MAP			
						True adjacencies reproduced	False adjacencies added	Total Area Deviation	Figure
	N	$\|E\|$	$\|E^c\|$	$\max \omega_r$	$\min \omega_r$	$\|E \cap E_{\mathbf{P}}\|$	$\|E^c \cap E_{\mathbf{P}}\|$	$\sum_{r=1}^{N} \|\omega_r - area(P_r)\|$	
`Blood`	8	19	9	0.374	0.006	14	2	0.152	4
						14	1	0.157	5
						15	0	0.165	6
`Netherlands`	12	22	44	0.212	0.023	17	7	0.239	7
						17	6	0.252	8
						18	5	0.259	9
						19	2	0.296	10
						20	1	0.390	11
						22	2	0.472	12
						22	1	0.509	13
`Germany`	16	28	92	0.198	0.001	22	10	0.151	14
						23	10	0.156	15
						24	7	0.329	16
						25	8	0.172	17
						25	4	0.365	18
						26	4	0.515	19
						27	6	0.537	20
						27	4	0.583	21
						27	3	0.781	22
						28	9	0.996	23

Table 1: Rectangular maps with different tradeoffs between an accurate representation of E by $E_{\mathbf{P}}$ versus an accurate representation of weights by areas

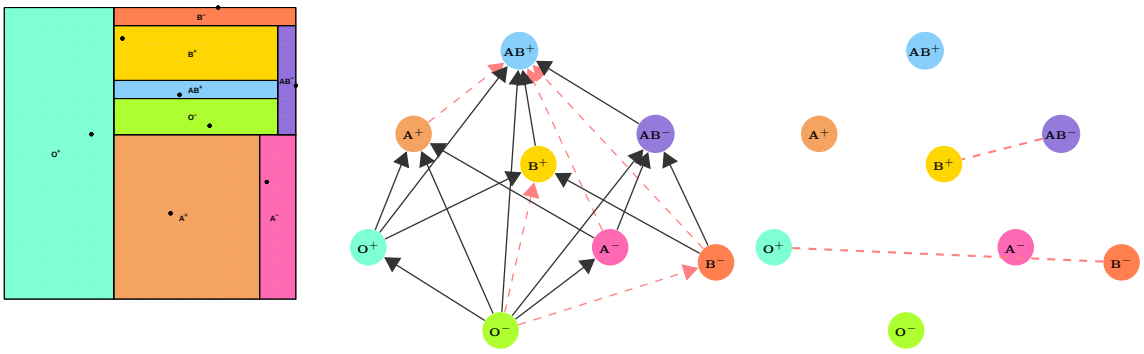


Figure 4: `Blood` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 14$, $|E^c \cap E_{\mathbf{P}}| = 2$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.152$.

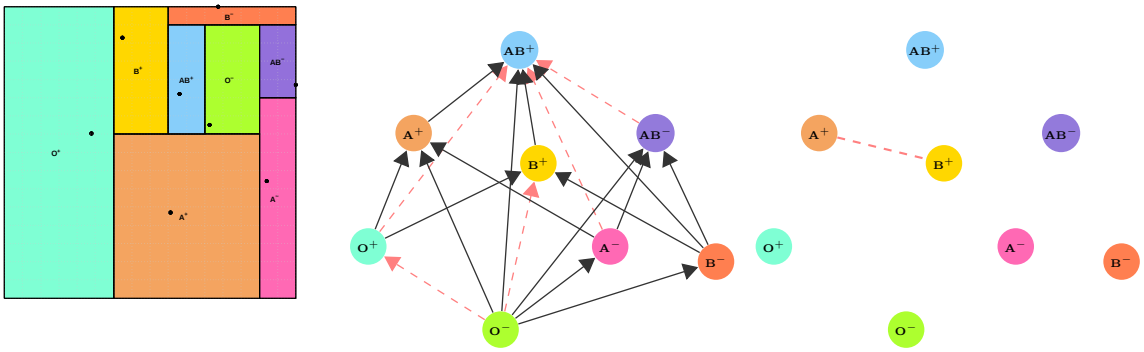


Figure 5: `Blood` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 14$, $|E^c \cap E_{\mathbf{P}}| = 1$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.157$.

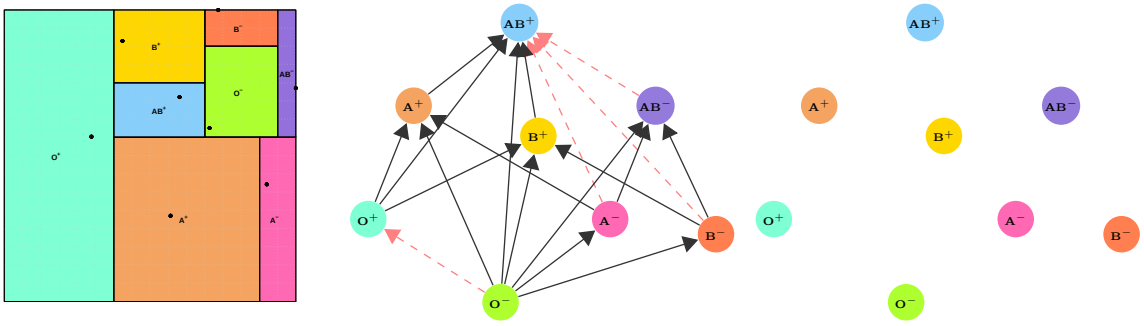


Figure 6: `Blood` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 15$, $|E^c \cap E_{\mathbf{P}}| = 0$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.165$.

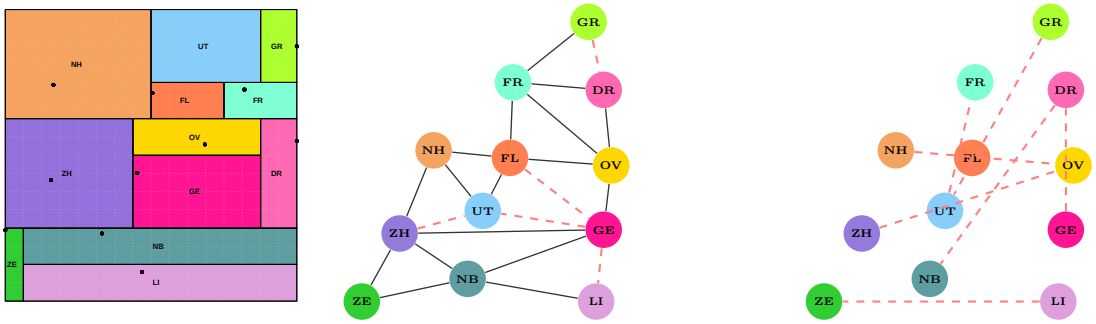


Figure 7: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 17$, $|E^c \cap E_{\mathbf{P}}| = 7$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.239$.

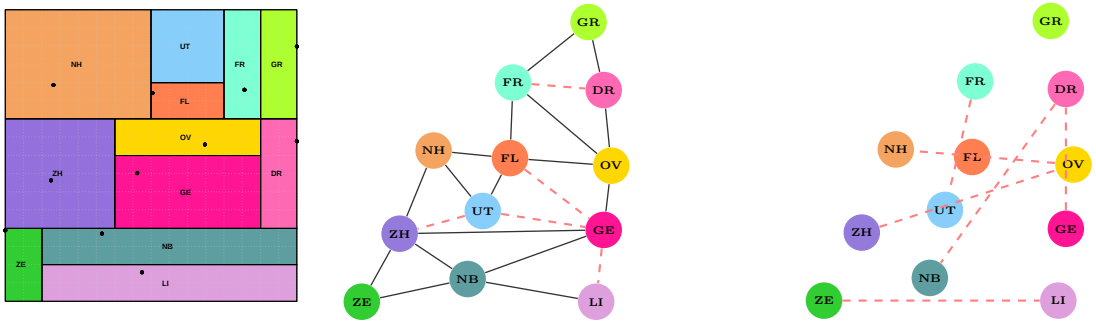


Figure 8: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 17$, $|E^c \cap E_{\mathbf{P}}| = 6$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.252$.

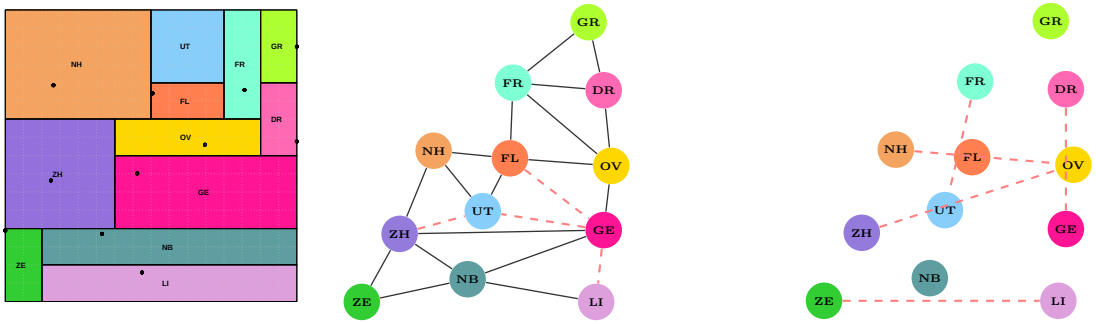


Figure 9: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 18$, $|E^c \cap E_{\mathbf{P}}| = 5$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.259$.

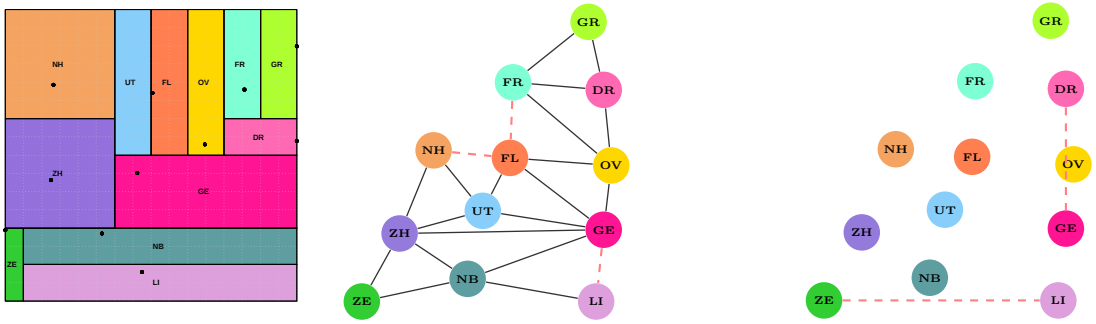


Figure 10: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 19$, $|E^c \cap E_{\mathbf{P}}| = 2$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.296$.

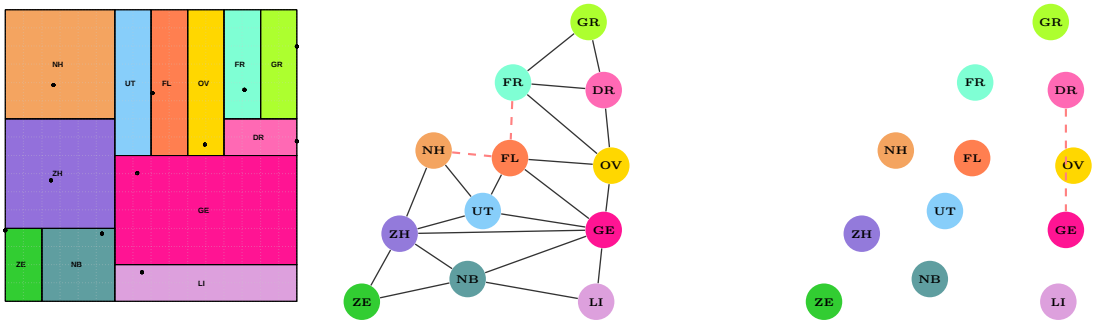


Figure 11: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 20$, $|E^c \cap E_{\mathbf{P}}| = 1$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.390$.

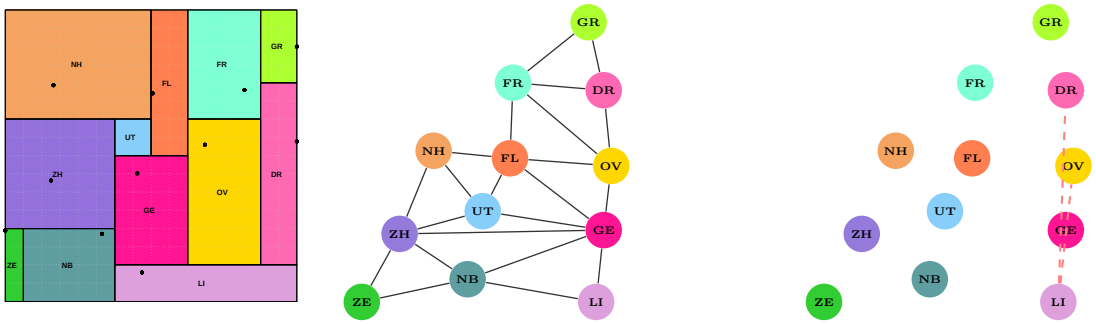


Figure 12: `Netherlands` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 22$, $|E^c \cap E_{\mathbf{P}}| = 2$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.472$.

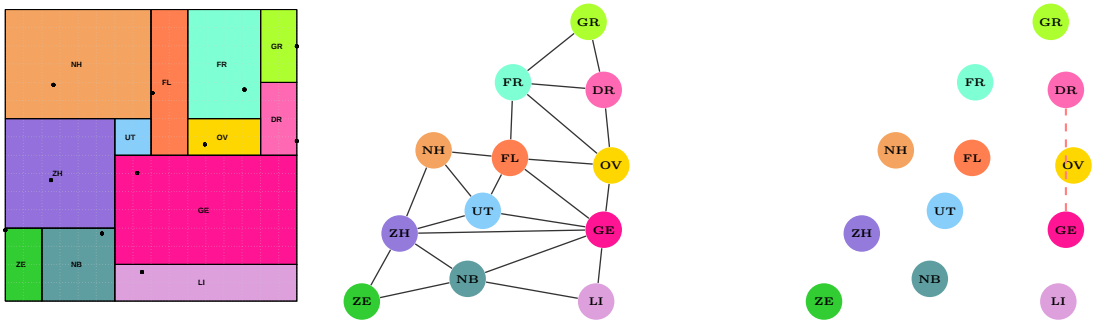


Figure 13: `Netherlands` rectangular map with objectives $|E \cap E_{\mathbf{P}}| = 22$, $|E^c \cap E_{\mathbf{P}}| = 1$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.509$.

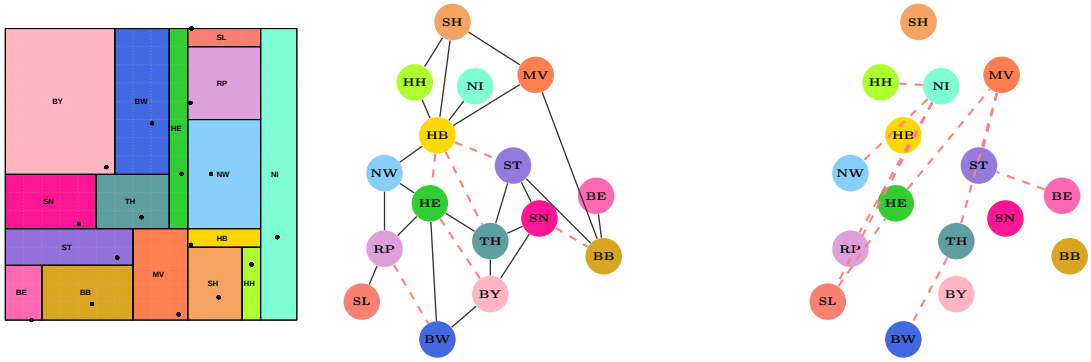


Figure 14: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 22$, $|E^c \cap E_{\mathbf{P}}| = 10$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.151$.

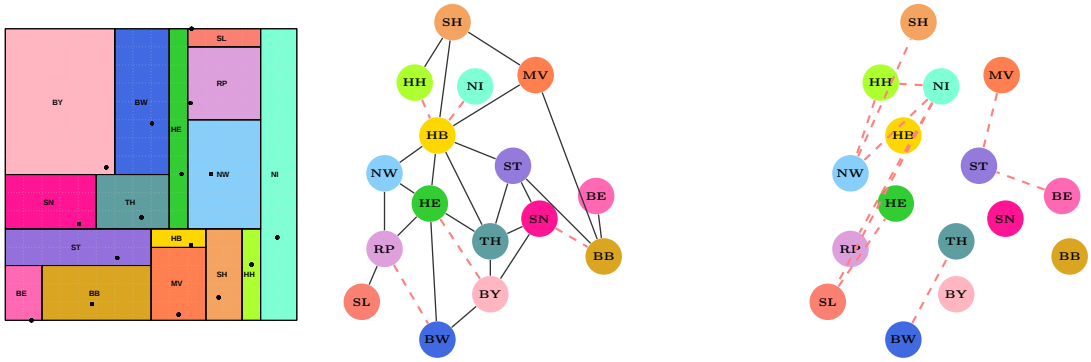


Figure 15: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 23$, $|E^c \cap E_{\mathbf{P}}| = 10$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.156$.

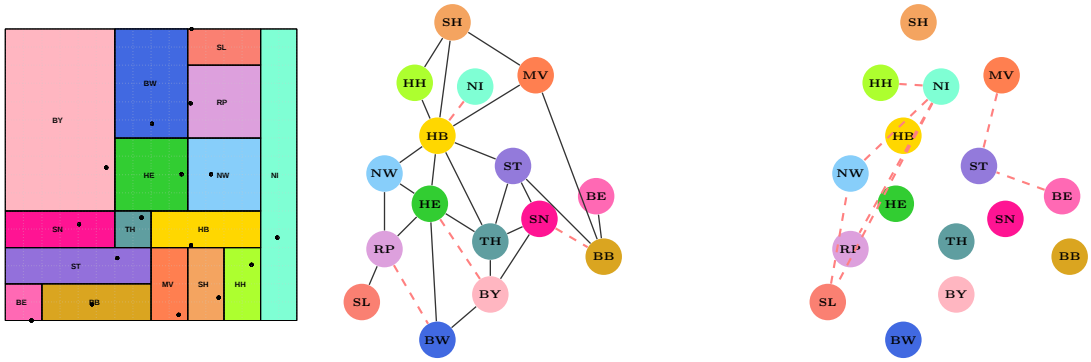


Figure 16: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 24$, $|E^c \cap E_{\mathbf{P}}| = 7$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.329$.

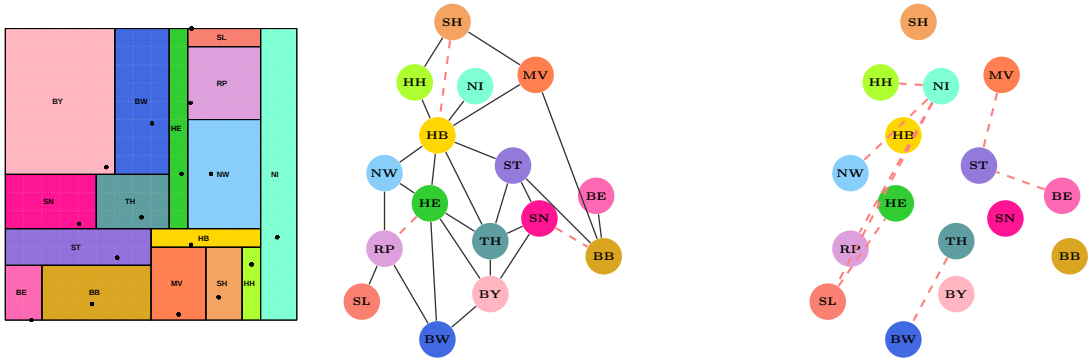


Figure 17: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 25$, $|E^c \cap E_{\mathbf{P}}| = 8$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.172$.

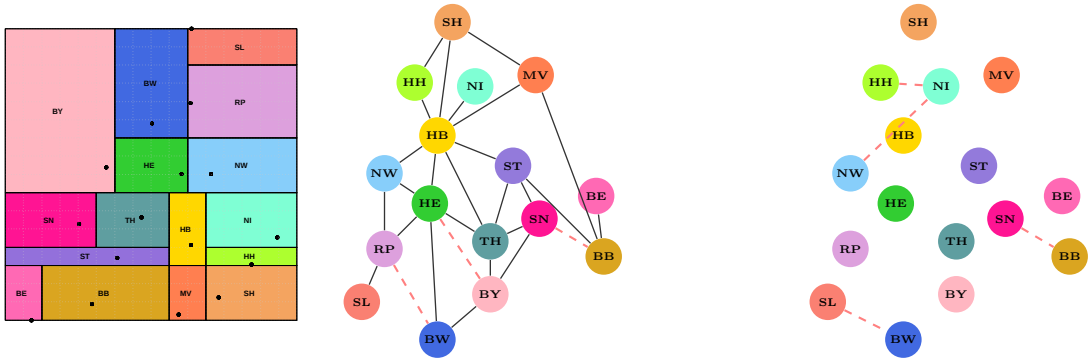


Figure 18: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 25$, $|E^c \cap E_{\mathbf{P}}| = 4$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.365$.

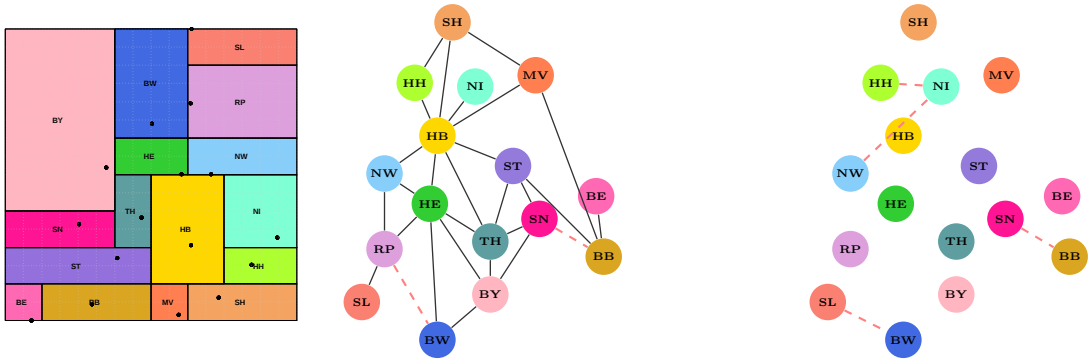


Figure 19: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 26$, $|E^c \cap E_{\mathbf{P}}| = 4$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.515$.

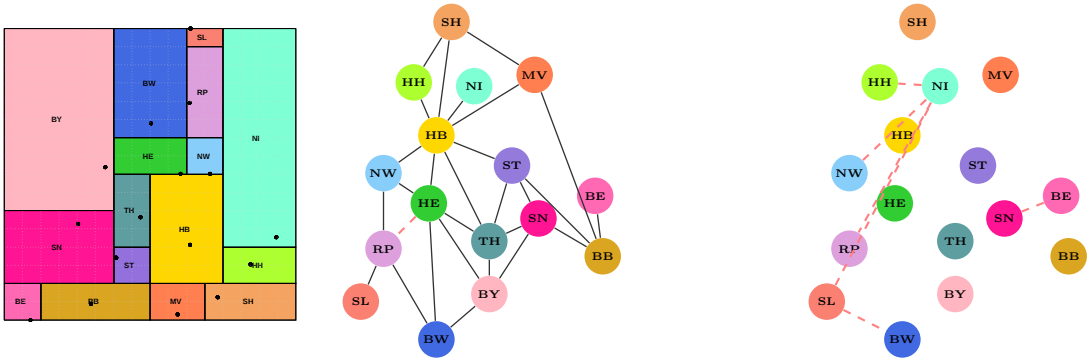


Figure 20: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 27$, $|E^c \cap E_{\mathbf{P}}| = 6$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.537$.

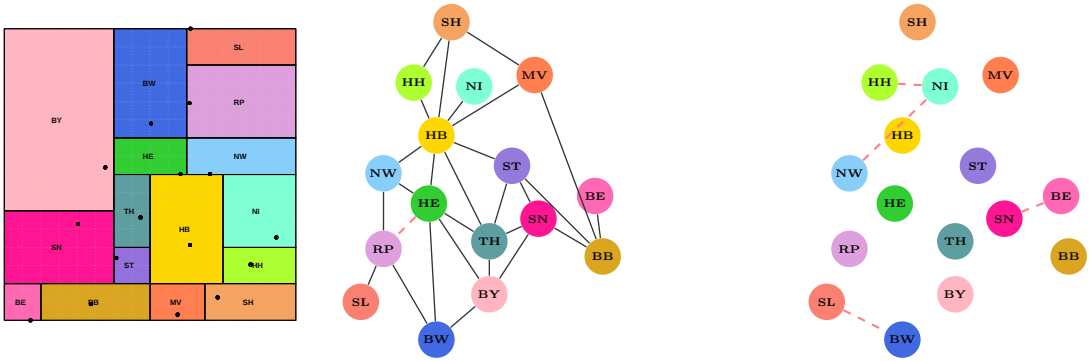


Figure 21: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 27$, $|E^c \cap E_{\mathbf{P}}| = 4$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.583$.

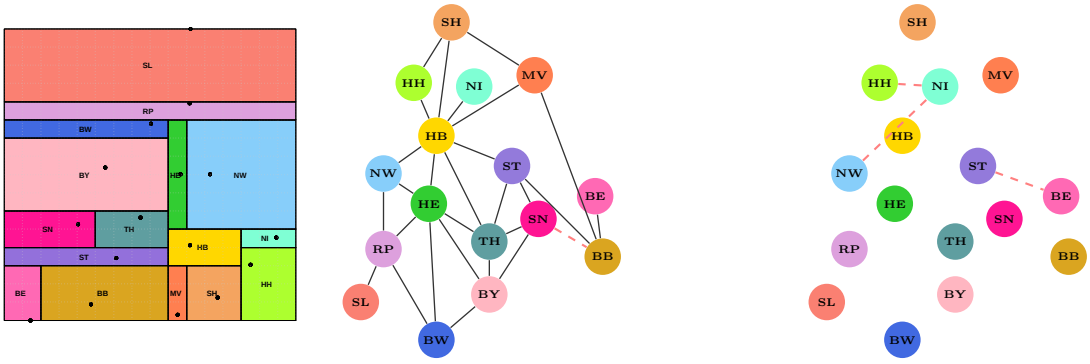


Figure 22: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 27$, $|E^c \cap E_{\mathbf{P}}| = 3$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.781$.

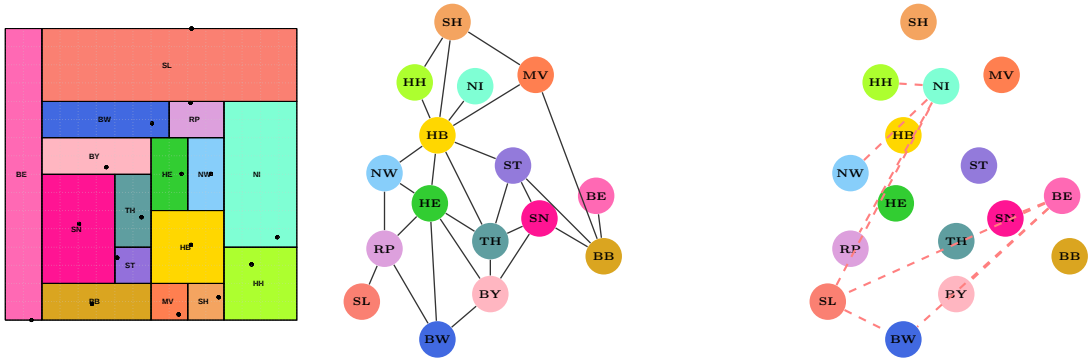


Figure 23: `Germany` rectangular map with objectives
$|E \cap E_{\mathbf{P}}| = 28$, $|E^c \cap E_{\mathbf{P}}| = 9$, $\sum_{r=1}^{N} |\omega_r - area(P_r)| = 0.996$.