

Chapter 23

Districting Problems

Jörg Kalcsics

Abstract Districting is the problem of grouping small geographic areas, called basic units, into larger geographic clusters, called districts, such that the latter are balanced, contiguous, and compact. Balance describes the desire for districts of equitable size, for example with respect to workload, sales potential, or number of eligible voters. A district is said to be geographically compact if it is somewhat round-shaped and undistorted. Typical examples for basic units are customers, streets, or zip code areas. Districting problems are motivated by quite different applications ranging from political districting over the design of districts for schools, social facilities, waste collection, or winter services, to sales and service territory design. Despite the considerable number of publications on districting problems, there is no consensus on which criteria are eligible and important and, moreover, on how to measure them appropriately. Thus, one aim of this chapter is to give a broad overview of typical criteria and restrictions that can be found in various districting applications as well as ways and means to quantify and model these criteria. In addition, an overview of the different areas of application for districting problems is given and the various solution approaches for districting problems that have been used are reviewed.

Keywords Districting criteria • Political districting • Sales territory design • Service districting

23.1 Introduction

Most problems discussed in this book focus on the location of facilities: where to locate, how many to locate, when to locate, which type to locate, etc. However, although the driving force is the location of facilities, equally important is the second aspect of location problems that is usually not mentioned explicitly: the allocation of customers to facilities. Even if this task is trivial in many classical location problems like the p -median or the p -center problem (see Chaps. 2 and 4), only after deciding

J. Kalcsics (✉)

Institute of Operations Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: kalcsics@kit.edu

about allocations can we evaluate a given facility configuration and, thus, try to find the optimal one. Hence, the allocations have a fundamental impact on the location of facilities and different rules of allocation will result in different evaluations of the same facility configuration. The aim of districting problems is now the other way around: we first find allocations—or, more generally, determine which customers should be served together—and then, if necessary, we find locations for the facilities serving the customers.

In general, districting is the problem of grouping small geographic areas, called basic units, into larger geographic clusters, called districts, in a way that the latter are acceptable according to relevant planning criteria. Typical examples for basic units are customers, streets, or zip code areas. Depending on the practical context, districting is also called territory design, territory alignment, zone design, or sector design. Three important criteria are balance, contiguity, and compactness. Balance describes the desire for districts of equitable size with respect to some performance measure for the districts. Depending on the context, this criterion can either be economically motivated, for example, equal sales potentials, workload, or number of customers, or have a demographic background, for example, the same number of inhabitants or eligible voters. A district is called contiguous if it is possible to travel between the basic units of the district without having to leave the district. Finally, a district is said to be geographically compact if it is somewhat round-shaped, undistorted, and without holes. Contiguous and compact districts usually reduce the travel time of the person responsible for servicing the district. Unfortunately, a rigid and concise mathematical definition of contiguity and compactness is often difficult and strongly depends on the available data. In addition, for each district often the location of a “facility” is either given or should be sought. This facility can be a branch office, a depot, or the home address of a sales person. Figure 23.1 shows an example of a districting plan for streets and for zip code areas.

Districting problems are motivated by quite different applications ranging from political districting over the design of districts for schools, social facilities, waste collection, or winter services, to sales and service territory design. Looking at the literature, it is striking that only few authors consider the districting problem independently from a practical background. Therefore, the aim of this chapter is to

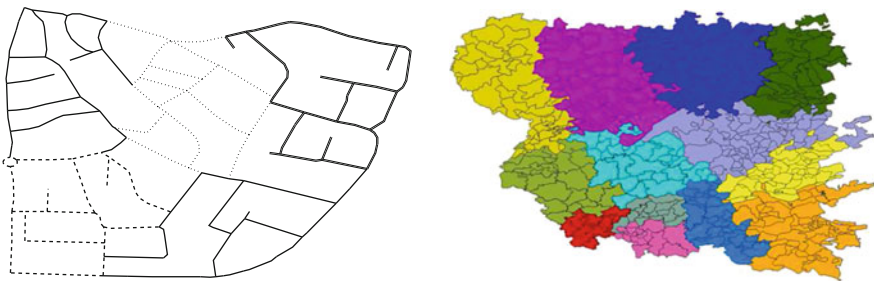


Fig. 23.1 An example of a districting plan for streets and for zip-code areas

give a broad overview of typical criteria and restrictions that can be found in the various districting applications as well as ways and means to quantify and model these criteria. As most districting applications have a strong spatial component, it is natural to integrate the algorithms into a Geographic Information System (GIS). In a modern GIS, users can access and utilize the rich variety of maps, spatial databases, and geographical objects available to appropriately mark out the problem and display the solutions, see also Chap. 19.

The rest of the chapter is organized as follows. The next section reviews the broad range of districting applications and identifies and motivates the different planning restrictions. In Sect. 23.3, basic notations are introduced. The next section discusses the most common criteria found in districting applications and discusses possible approaches to quantify these criteria and to incorporate them into districting models. Finally, Sect. 23.5 presents an overview of the different solution techniques for solving districting problems.

23.2 Applications

There are four major areas of application for districting problems: political districting, sales territory design, service districting, and distribution districting, and this section provides a comprehensive but non-exhaustive overview. But before we start, we mention a first “application” in the context of facility location that derives from the problem of aggregating demand points for location problems with the aim of reducing the complexity of the problem. Simchi-Levi et al. (2003) formulate the following guidelines (among others): aggregate demand points for 150–200 zones, make sure each zone has an approximately equal amount of total demand, and place aggregated points at the center of the zone. These guidelines read as a classical districting problem.

23.2.1 Political Districting

Political districting is the problem of dividing a governmental area, such as a city or a state, into constituencies from which political candidates are elected. Basic units typically correspond to census tracts, which are given as polygons, and the districts to the electoral constituencies. In general, the process of redistricting has to be periodically undertaken to account for population shifts. The length of these periods varies from country to country, e.g., in New Zealand every 5 years, in Canada and the U.S. every decade (after each census). In the past, political districting has often been flawed by manipulation aiming to favor some particular party or to discriminate against social or ethnic minorities. Since the responsibility for approving state and local districting plans usually falls to elected representatives, plans are likely to be shaped implicitly, if not overly, by political considerations, e.g., to keep them in

power. A famous case arose in Massachusetts in the early nineteenth century when the state legislature proposed a salamander-shaped electoral district in order to gain electoral advantage. The governor of the state at that time was Elbridge Gerry, and this practice became known as gerrymandering. See Lewyn (1993) for an interesting description of gerrymandering cases.

To avoid political interference, many states have set up a neutral commission to determine political boundaries satisfying a number of legislative and common sense criteria. Depending on the country or jurisdiction involved, these criteria may be enforced by legislative directive, judicial mandate, or historical precedent. However, there is no consensus in political science, law, or geography on which criteria are legitimate for the districting process, i.e., satisfy the neutrality condition. Moreover, it is often unclear how they should be measured (Williams 1995). One important issue at stake is population equality. To respect the principle of “one man-one vote”, i.e., every vote has the same power, all districts should contain approximately the same number of voters, i.e., be balanced. In the U.S., population equality has been deemed by the courts to be very important, and as a result, the total deviation of congressional districts from perfect balance was less than 1 % after the last census in 2000 (Webster 2013). In other countries, the allowed deviations are usually higher (Handley and Grofmann 2008). Two other important criteria always being mentioned are contiguity and compactness which both aim at preventing gerrymandering. While contiguity is generally undisputed and easy to verify, this is not the case for compactness. There is a broad discussion on how to quantify this criterion adequately (Horn et al. 1993), and whether it is relevant in the first place because an algorithm will never gerrymander on purpose as long as it does not use political data (Garfinkel and Nemhauser 1970). Moreover, if an adequate minority representation is sought for, this may sometimes only be achieved through non-compact districts (Dixon 1968). Other—often disputed—criteria are the conformity to administrative boundaries, e.g., cities or counties, the preservation of communities of interest, socio-economic homogeneity or a fair representation of minority voters across the districts, the similarity with the previous electoral districts, or the consideration of topological obstacles, like mountain ranges, lakes, or rivers (cf. George et al. 1997; Parker 1990; Bozkaya et al. 2011). An excellent review on typical criteria for political districting and their eligibility is given in Webster (2013).

When discussing automated procedures in the literature, it is always noted that they are non-partisan and neutral as long as they do not use political data and, hence, prevent gerrymandering. However, even if the computer does not gerrymander on purpose, it may still do it accidentally, precisely because no political data is taken into account. Therefore, Puppe and Tasnádi (2008) recently introduced the notion of an (ex post) unbiased districting plan. In such a plan the number of districts won by each party respects the relative strength of the party in the population as close as possible. They focus on game theoretical aspects of the problem; see also Nagel (1965). However, one has to do a careful weighing up to avoid forthright politically biased criteria that lead, in spirit, to gerrymandering.

23.2.2 *Sales Territory Design*

The important but expensive task of designing sales territories is common to all companies that operate a sales force and need to subdivide the market area into regions of responsibility that are each attended to by one or more sales representatives. According to Zoltners and Sinha (2005), approximately every tenth full-time employee in the U.S. is working as a field and retail sales person and the expenditure for them is more than three trillion dollars every year. Territories with low sales potential, intense competition, or too many small accounts lead to low morale, poor performance, a high turnover rate, and an inability to assess the productivity of individual territories. Therefore, well-planned decisions enable an efficient market penetration and lead to decreased costs and improved customer service and sales. Zoltners and Sinha (2005) “guestimate” that a good territory alignment can increase sales by 2–7% compared to an average alignment. In the related literature, districts are predominantly called territories and districting is termed territory alignment or territory design.

In the classical problem, the task is to assign a given set of (prospective) customer accounts, each with a fixed market potential, to the individual members of the sales force such that each customer has a unique representative and each sales person faces equitable workload and travel time and has an equal income opportunity in terms of incentive pay (Zoltners and Sinha 2005). Thus, basic units correspond to accounts and are usually given as points. Concerning the travel time, if a sales person visits each customer every day, then the travel time is proportional to the length of a TSP tour. However, the workload of districts is usually balanced over 2–4 weeks and some customers may have to be visited only once during this time whereas others require weekly service. Moreover, customers may have time windows, tours may include overnight stays, and so on, which makes the actual computation of the travel times almost impossible. Hence, in most cases one has to rely on estimates. Typically, a sales person is exclusively responsible for all customers within a specific geographic region. However, in large companies sometimes a sales person is only responsible for a certain product segment or accounts of a particular size within his region. In such cases, sales territories may overlap. For practical examples of sales territory design see Fleischmann and Paraschis (1988), Zoltners and Sinha (2005), López-Pérez and Ríos-Mercado (2013).

Three classical sales districting criteria are again balance, contiguity, and compactness. In contrast to political districting, typically more than one performance measure has to be balanced, for example workload and sales potential. A district with comparatively many small accounts or customers with low sales potential will yield lower sales and, hence, lower incentives for the responsible sales person than a district with an equitable workload but only high potential accounts. This disparity will lead to discontent among the sales persons and, in the long run, lower sales for the company. Having said that, only few authors consider more than one balancing criterion: Deckro (1977), Zoltners and Sinha (1983), Ríos-Mercado and Fernández (2009). Contiguous districts are desired to obtain clearly defined geographic areas

of responsibility and, together with compactness, to reduce the unproductive travel time of the sales force. Unfortunately, as basic units are typically points, it is not clear how to assess contiguity. Moreover, it is important to point out that the desire for compact districts is born out of necessity because the actual travel times are usually impossible to determine efficiently. The hope is that geographically compact and contiguous districts result in smaller travel times on a day-to-day basis than non-compact and/or non-contiguous districts.

As the main goal of most companies is to maximize profit, several authors relax the assumption that the sales potential of customers is fixed. Instead, they propose an integration of time-effort allocation and territory design methods to increase profit while maintaining the equitable workload criterion (cf. Lodish 1975; Glaze and Weinberg 1979; Zoltners and Sinha 1983). These models not only assign customers to sales people but also determine how much time should be invested in the customer. Some authors even object that equity is not the primary goal for most companies. Instead, the aim should be to maximize profits, regardless of any balancing aspect (Skiera and Albers 1994; Drexl and Haase 1999). However, in practice sales persons are typically reluctant to implement such detailed call plans resulting from pure profit maximizing approaches (Zoltners and Sinha 2005). Moreover, designing territories is a mid- or even long-term decision whereas time-effort allocation is an operational problem that is influenced by weather (espc. in the beverage industry), sales promotions, etc. Thus, these two problems should be addressed separately.

Often, the number of districts to be designed is predetermined by the designated sales force size (Fleischmann and Paraschis 1988). If the size is not self-evident, methods based on the total workload involved in covering the entire market compared to the available time per sales person can be used. Another possibility is to follow the “decreasing returns” principle and add sales persons to the sales force as long as the expected increase in profit exceeds the expected increase in costs (Howick and Pidd 1990; Zoltners and Sinha 2005).

As sales persons have to visit their territories regularly, their home-base, e.g., office or residence, is an important factor to be considered in the alignment process. However, there is no consensus as to whether predetermined locations should be kept or be subject to the planning process. On the one hand, most sales persons have strong preferences for home-base cities. Hence, such locations should be respected or determined prior to the alignment to socialize them with the sales management (Zoltners and Sinha 2005). On the other hand, addresses and sales personnel frequently change and the management often does not want sales persons residences to overly influence the definition of territories (Fleischmann and Paraschis 1988).

23.2.3 Service Districting

The problem of designing service districts appears in various contexts. One area of applications focuses on social facilities, like hospitals or public utilities. Sometimes

districts are needed to define for each inhabitant which facility he should visit to obtain service, for example for preventive medical examinations, or to determine areas of responsibility of home-care visits by healthcare personnel, like nurses or physiotherapists. The goal is to determine contiguous districts that have a good accessibility with respect to public transportation and have an equitable workload based on service and travel time or account for a high capacity utilization of the social facility (cf. Minciardi et al. 1981; Blais et al. 2003; Benzarti et al. 2013).

A second field of applications deals with providing service to streets. A classical problem concerns the design of districts for postal or leaflet delivery. Instead of considering each household separately, districts are composed of whole streets. Thus, basic units correspond to streets and each basic unit typically has two attributes: the times required to traverse the street with and without providing service. The task is to partition the streets into a given number of districts such that the required delivery time is approximately the same for all districts and does not exceed the working time restriction of the deliverer. The delivery time is proportional to the length of a Chinese postman tour through the district, which can be computed efficiently. Moreover, the delivery districts should be contiguous, incur little deadheading, and should not overlap, i.e., be geographically compact (Bodin and Levy 1991; Butsch et al. 2014). A common characteristic of these applications is that the deliverer either walks through his district on foot or goes by bike so that one-way streets are no hindrance. If a street is too wide or has too much traffic to serve it in a zig-zag pattern, then each side of the street is modeled as a separate basic unit. A similar problem arises in the context of meter reading in power distribution networks (Silva de Assis et al. 2014). Closely related are districting problems for solid waste disposal, salt spreading, and winter gritting (Hanafi et al. 1999; Muyltermans et al. 2002; Lin and Kao 2008). The criteria are almost identical to postal delivery. The only differences are that vehicles typically have to respect one-way streets and have difficulties making U-turns, and that their tours have to include a depot, e.g., to drop off waste or refill salt. All these aspects make the computation of the travel times more difficult. Other applications deal with the design of patrol districts for police cars and primary response areas for ambulances, where the districts additionally should have an average response time and/or incident arrival rate below a given threshold (Baker et al. 1989; D'Amico et al. 2002; Xu and Yum 2010).

Other applications deal with the problem of assigning residential areas to schools (Ferland and Guénette 1990; Schoepfle and Church 1991). Criteria to be taken into account are capacity limitations and an equal utilization of the schools, maximal or average travel distances for students, good accessibility, and ethnic balance. Another aspect is to decide which students should walk to school and which should take the school bus. Districting problems also occur in electric power networks. According to Bergey et al. (2003), the World Bank regularly faces the challenge of helping developing countries to move from state owned, monopolistic electric utilities to a more competitive environment with multiple electricity service providers. At that, they face the task of partitioning the physical power grid into economically viable districts (distribution companies). The main aim is to determine non-overlapping and contiguous districts with approximately

equal revenue potential (to foster competition) which are compact over a geographic region (to be easier to manage and more economical to maintain).

23.2.4 *Distribution Districting*

Another important field of applications is the design of pickup and delivery districts in logistics. Typically, such problems are modeled and solved as vehicle routing problems. However, if there exists considerable uncertainty in the demand of customers, several authors propose a two-phase approach that first builds the pickup and delivery districts and then does the routing on a day-to-day basis. This conforms with the well-known “cluster first–route second” paradigm for vehicle routing problems. Hence, basic units correspond to potential customers, given as points, and the task is to partition the set of customers into districts, one for each driver, such that the districts satisfy certain planning criteria. A first advantage of these fixed customer assignments is that the driver becomes familiar with his district. This, in turn, increases the driver’s performance since he becomes quicker at finding customer addresses, localizing offices within buildings as well as organizing his routes (Zhong et al. 2007). A second advantage is that customers become familiar with their drivers, which increases customer satisfaction (Jarrah and Bard 2012). These advantages however have to be carefully weighed against flexible customer assignments on a daily basis which enable the planner to maximize the driver utilization and minimize the routing costs (Zhong et al. 2007).

Concerning the criteria for the districting process, districts should be contiguous and compact, and the workload should either be balanced or at least not exceed a given upper bound, e.g., the driver working time. The workload includes the service time at the customers and typically also an estimate of the average travel time within the district and to a centralized depot (Galvão et al. 2006; Haugland et al. 2007; Zhong et al. 2007; Jarrah and Bard 2012; Lei et al. 2012).

A final application concerns the establishment of a distribution center which involves a considerable level of risk due to its enormous start-up investment and volatile customer demand patterns. One way of reducing this risk is to avoid both overcrowding and, especially, underutilization of centers by balancing the allocation of customers to them (Zhou et al. 2002).

23.3 Notations

This section introduces notations for the main components of districting problems.

23.3.1 Basic Units

A districting problem comprises a set $J = \{1, \dots, n\}$ of *basic units*, sometimes called sales coverage units, basic areas, or geographical units. Each basic unit represents a geometric object in the plane: a point, e.g., a geo-coded address, a line segment, e.g., a street, or a polygonal area, e.g., a zip code area, county, or predefined company trading area. The distance between two basic units $i, j \in J$ is denoted as $d_{ij} = d(i, j)$. Typical examples for d_{ij} are Euclidean (cf. Fleischmann and Paraschis 1988) or road distances (cf. Ríos-Mercado and Salazar-Acosta 2011). The latter have the advantage that they can properly reflect obstacles like rivers or mountain ranges. For non-point objects, distances are either computed between representative points, e.g., the midpoint of a street or the centroid of a polygon, or as the surface-to-surface distance.

Moreover, one or more quantifiable attributes, called *activity measures*, are associated with each basic unit. Typical examples are service times, estimated sales potential, or number of voters. Sometimes, they also include an estimate of the travel time for visiting the basic unit (Jarrah and Bard 2012). The activity measures are all assumed to be deterministic. Let w_j^q denote the q -th activity measure of basic unit $j \in J$, $1 \leq q \leq Q$, where Q is the number of different attributes to be considered. If $Q = 1$, the superscript is usually omitted.

If explicit neighborhood information is given for the basic units, then $G = (V, E)$ denotes the *neighborhood* or *contiguity graph* where $v_j \in V$ corresponds to $j \in J$ and $\{v_i, v_j\} \in E$ iff basic units i and j are neighboring. The length of edge $\{v_i, v_j\}$ is d_{ij} . Finally, $N(j) \subseteq V$ denotes the set of basic units adjacent to $v_j \in V$.

23.3.2 Districts

A *district* D_k , $1 \leq k \leq p$, is a subset of basic units, where p is the total number of districts. The number of districts can either be fixed in advance, e.g., the number of political districts to create or the number of available nurses for elderly care, or be subject to planning, e.g., the minimal number of salespersons required to service all customers or the minimal number of patrol cars to ensure a certain response time. The q -th activity measure of a district is the sum of the activity measures of its basic units, i.e., $w^q(D_k) = \sum_{j \in D_k} w_j^q$. For $Q = 1$, $w^1(D_k)$ is simply called the *size* of the district. Note that sometimes the size also includes an estimate of the (expected) travel time. However, as travel times are represented through the compactness criterion, we refrain from including them and just mention when this may change things.

In some applications the location c_k of a facility is associated with each district D_k . This may be some predefined site, e.g., a hospital providing preventive medical care, or be an outcome of the districting process, e.g., the optimal location of a sales

office. In districting, this location is called the *center* of the district. One has to be aware of the ambiguity with the notion of a center in location theory, which is something different, see Chap. 4. Typically, the center coincides with a basic unit, i.e., $c_k \in J$. A predetermined set of centers is denoted by J_c .

Finally, a *districting plan* \mathcal{D} is defined as a set of p districts $\mathcal{D} = \{D_1, \dots, D_p\}$.

23.3.3 Problem Formulation

The districting problem can now informally be described as follows: Partition all basic units J into a number of p districts that satisfy the planning criteria of balance, compactness, and contiguity and, if required, locate a center within each district. Unfortunately, in contrast to many other optimization problems, there does not exist *the* mathematical model for districting problems. This is mainly due to the considerable ambiguity on how to quantify the different planning criteria and in the motivation and relevance of some of them.

23.4 Districting Criteria

This section presents an overview over typical criteria employed in districting problems and various ways and means to quantify them. In the following, a measure for a criterion applied to a single district (the whole districting plan) is termed a local (global) measure. Moreover, if not explicitly stated otherwise, let $Q = 1$.

23.4.1 Complete and Exclusive Assignment

In most cases, each basic unit is assigned to exactly one district, i.e., the districts define a partition of the set J of basic units:

$$D_1 \cup \dots \cup D_p = J \quad \text{and} \quad D_l \cap D_k = \emptyset, \quad 1 \leq l, k \leq p, l \neq k.$$

The requirement of exclusive assignment is sometimes also termed *integrity*. For political districting, these criteria are obvious. In sales territory design, unique allocations result in transparent responsibilities for the sales force avoiding contentions and allowing the establishment of long-term customer relations.

23.4.2 Balance

This criterion is one of the trademarks of districting problems. It expresses the desire for districts of equitable size with respect to the activity measure(s). In political districting, this criterion is employed to ensure the “one man–one vote” principle, and in sales territory design to avoid districting plans with large discrepancies in terms of workload, sales potential, or travel time.

Due to the discrete structure of the problem and the integrity assumption, perfectly balanced districts can generally not be accomplished. There exist different approaches in the literature to quantify imbalance and to incorporate the criterion into the districting process. The most common local measure is based on the relative deviation of the district size $w(D_k)$ from the mean district size $\mu = w(J)/p$:

$$bal(D_k) = \left| \frac{w(D_k) - \mu}{\mu} \right|, \quad 1 \leq k \leq p$$

(cf. Forman and Yue 2003; Ríos-Mercado and Fernández 2009; Silva de Assis et al. 2014). The larger this deviation is, the worse is the balance. A district D_k is perfectly balanced, if $bal(D_k) = 0$. If the district sizes also contain a solution dependent performance measure, like travel times, then this affects μ and the balance of one and the same district may be different in different districting plans. Another approach concedes a priori a certain relative deviation $\alpha > 0$ from perfect balance and only measures the imbalance exceeding this threshold (Bodin and Levy 1991; Bozkaya et al. 2011)

$$bal(D_k) = \frac{1}{\mu} \max\{w(D_k) - (1 + \alpha)\mu, (1 - \alpha)\mu - w(D_k), 0\},$$

i.e., the district is balanced if its size is between this lower and upper bound. Instead of determining the bounds based on the mean district size, they are sometimes directly motivated by the application, e.g., the working time restrictions of the mailman or the sales potential required to ensure a decent living for the sales person.

Using these local measures, the global balance of a districting plan is then typically computed as the maximal balance of a district

$$bal^{max}(\mathcal{D}) = \max_{k=1, \dots, p} bal(D_k).$$

Less common are the sum over all districts (Bozkaya et al. 2003; Bodin and Levy 1991) or a convex combination of both (Butsch et al. 2014):

$$bal^{sum}(\mathcal{D}) = \sum_{k=1}^p bal(D_k) \quad \text{and} \quad bal^{cv}(\mathcal{D}) = \lambda bal^{sum}(\mathcal{D}) + (1 - \lambda) bal^{max}(\mathcal{D}),$$

with $\lambda \in (0, 1)$. The convex combination alleviates some of the weaknesses of bal^{sum} and bal^{max} . The latter does not take into account the balance of all districts and sometimes yields rather poor solutions on average whereas the former allows a few highly unbalanced districts to be compensated by some well-balanced districts. A different global approach computes the range of district sizes (Tavares-Pereira et al. 2007)

$$bal^{mg}(\mathcal{D}) = \max_{k=1,\dots,p} w(D_k) - \min_{k=1,\dots,p} w(D_k).$$

23.4.2.1 Mathematical Modelling

In districting models, there is no clear trend on whether to treat balance as a hard constraint (Hess et al. 1965; Fleischmann and Paraschis 1988; Zoltner and Sinha 2005) or to include it in the objective function (Blais et al. 2003; Ricca and Simeone 2008; Silva de Assis et al. 2014). In the former case, the size of each district is required to lie between a given lower and upper bound. Some authors even do both (Bergey et al. 2003; Salazar-Aguilar et al. 2013b). All of the above measures easily give rise to linear expressions.

23.4.3 Contiguity

Almost all districting approaches require districts to be contiguous. In political districting, this criterion should prevent gerrymandering. For the other types of applications, contiguous districts reduce the day-to-day travel distances for sales persons, delivery vans, snow ploughs, mailmen, etc. Unfortunately, a rigid and concise mathematical formulation of contiguity is difficult for basic units representing points.

23.4.3.1 Graph-Based Measures

If basic units are lines or polygons, it is easy to derive explicit neighborhood information. For example, two zip-code areas are neighboring if they share a common border, or two streets if they meet in a crossroad. In the former case, sometimes an additional requirement is the existence of a direct road connection between the two basic units. In general, two basic units are called *neighboring*, if their geometric representations have a nonempty intersection. This information is stored in the neighborhood graph $G = (V, E)$, and a district is contiguous if the basic units of the district induce a connected subgraph in G .

If basic units are represented by points, e.g., customer addresses, it is not clear how to assess contiguity. Over the years, different surrogate definitions for

contiguity have been proposed. One approach is based on proximity graphs to estimate the adjacency of points. One such graph is the Gabriel graph, in which two nodes v_i and v_j are connected by an edge if and only if the disc with antipodal points v_i and v_j does not contain any other node in its interior (Gross and Yellen 2003). A second approach to construct a contiguity graph is based on the Voronoi diagram (Lei et al. 2012). Two basic units are defined to be adjacent, iff their Voronoi cells have a common link within the smallest axis-parallel rectangle enclosing all basic units (for a definition of Voronoi diagrams and cells see Aurenhammar et al. 2013). A third construction of the proximity graph is to start with a complete graph and then sequentially go over all edges and delete for two intersecting edges in the planar representation of the graph the longer or more costly one (Haugland et al. 2007). All three graphs are planar. Moreover, by definition the Gabriel graph is a subset of the Voronoi-based graph.

Example 23.1 An example for these three proximity graphs for a point set with 26 basic units is depicted in Fig. 23.2. The Gabriel graph defines the most strict neighborhood relation. The graphs obtained by Lei et al. (2012) and Haugland et al. (2007) are fairly similar. The main difference is that the latter typically establishes more adjacencies along the boundary of the convex hull of the point set. Just by looking at the graphs it is difficult to decide which one is more suitable.

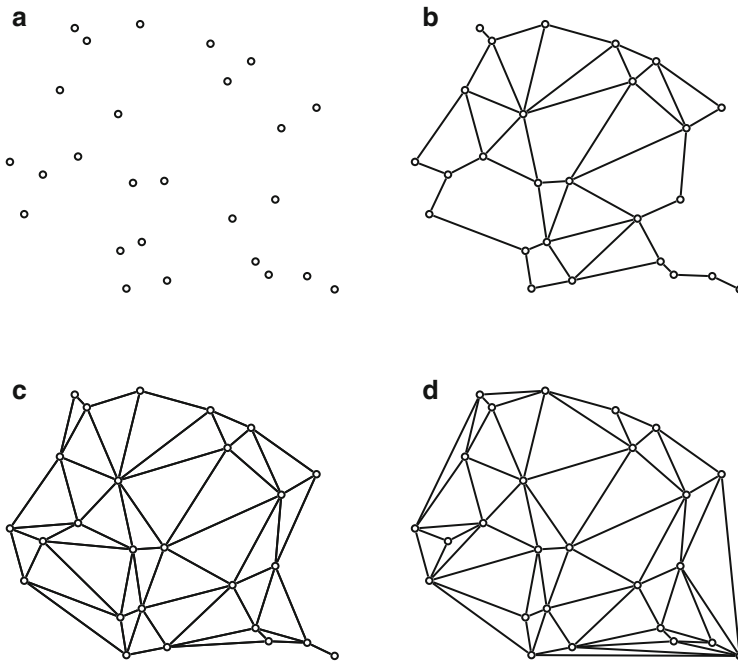


Fig. 23.2 Three different approximate contiguity graphs. (a) Point set of basic units. (b) Gabriel graph. (c) Voronoi-based graph. (d) Non-crossing edges graph

Finally, if the underlying road network is given, yet another possibility is to define two basic units as being adjacent, if the shortest path between the two does not contain another basic unit.

23.4.3.2 Geometric Measures

If no neighborhood information for basic units is given or can reasonably be derived, an alternative is to determine the overlap between the districts. For example, by computing the convex hull $ch(D_k)$ around each district D_k and defining a district to be contiguous if no basic unit of another district lies in its convex hull, i.e., $ch(D_k) \cap ch(D_l) = \emptyset, \forall l \neq k$ (Kalcsics et al. 2005; Jarrah and Bard 2012). One advantage of this approach is that convex districts usually prevent the crossing of routes of different districts, a characteristic that typically implies inefficient routes.

23.4.3.3 Mathematical Modelling

In districting models, contiguity is always treated as a hard constraint (except in Hanafi et al. 1999). One possibility to include it in a mathematical programming formulation is the following: Let $c_k \in J_c$ be the predetermined center of district k and $S \subseteq J \setminus \{N(c_k) \cup \{c_k\}\}$ be a subset of basic units that are not adjacent to c_k . If all elements of S are assigned to k , i.e., $S \subset D_k$, then at least one basic unit not in S that is adjacent to an element of S must also be assigned to k :

$$\sum_{j \in \bigcup_{i \in S} N(i) \setminus S} x_{kj} - \sum_{j \in S} x_{kj} \geq 1 - |S| \quad \forall S \subseteq J \setminus \{N(c_k) \cup \{c_k\}\},$$

where x_{kj} is 1 if $j \in J$ is assigned to district k and 0 otherwise (Drexler and Haase 1999). The drawback of this formulation is, that it requires an exponential number of constraints (although it gives naturally rise to a cut generation approach, Ríos-Mercado and López-Pérez 2013). A second possibility that only needs a linear number of constraints is based on network flow constraints. Each basic unit has one unit of supply, and the district centers act as sinks. District k is contiguous iff there exists a flow from its basic units to c_k that only passes basic units in D_k :

$$\begin{aligned} \sum_{i \in N(j)} f_{ji} - \sum_{i \in N(j)} f_{ij} &= x_{kj} & \forall j \in J \setminus \{c_k\} \\ \sum_{i \in N(j)} f_{ij} &\leq (n - 2) x_{kj} & \forall j \in J \setminus \{c_k\} \\ \sum_{i \in N(c_k)} f_{i,c_k} &\leq n - 1, \end{aligned}$$

where f_{ij} is the flow from basic unit i to j and $f_{c_k,j} = 0, \forall j \in N(c_k)$ (Shirabe 2009).

A simpler approach is to require that each district is a subtree of a shortest path tree $T(c_k)$ rooted at the district center c_k , where the edge lengths typically correspond to road distances or are all assumed to be 1. Then, for each basic unit j of district k , at least one of the adjacent basic units $i \in N(j)$ that immediately precedes j on some shortest path to the center c_k also has to be included in the district:

$$x_{kj} \leq \sum_{i \in S_j} x_{ki} \quad \forall j \in J \setminus \{c_k\},$$

where $S_j = \{i \in N(j) \mid i \text{ immediately precedes } j \text{ on some shortest path from } j \text{ to } c_k\}$ (Zoltners and Sinha 1983; Mehrotra et al. 1998). Although this excludes some contiguous districts, these are unlikely to be compact, as they typically have large protrusions or indentations, or contain enclaves.

It is straight forward to extend all of the above constraints to the case where the choice of district centers is part of the optimization. For geometric contiguity measures obviously only informal mathematical formulations can be derived.

Remark 23.1 Only few authors try to derive approximate neighborhood graphs for point-like basic units. The majority simply does not consider contiguity at all and tries to obtain districts with little overlap through an appropriate compactness measure, see also Example 23.3.

23.4.4 Compactness

A district is said to be geographically compact if it is somewhat round-shaped and undistorted. The motivation for compact districts is almost identical to ensuring contiguity: to prevent gerrymandering or to reduce the day-to-day travel distances within the districts. Although being a very intuitive concept, a rigorous definition of compactness does not exist and, moreover, strongly depends on the geometric representation of basic units. In the context of political districting, typically measures based on the shape of districts are employed whereas in sales and distribution districting, distance-based measures are predominant. In the following, the most common ones for both approaches are presented.

23.4.4.1 Geometric Measures

If basic units are given as polygons, geometric approaches based on the area or perimeter of a district can be used to quantify compactness. Two common local measures are the Reock and Schwartzberg tests. The former calculates the ratio

of the district area to the area of the smallest enclosing circle, while the latter determines the ratio of the districts perimeter length to the circumference of a circle with equal area

$$cmp(D_k) = \frac{A(D_k)}{\pi r_{enc}^2} \quad \text{and} \quad cmp(D_k) = \frac{P(D_k)}{2\sqrt{\pi A(D_k)}},$$

where $A(\cdot)$ and $P(\cdot)$ denote the area and the length of the perimeter, respectively, of a district and r_{enc} the radius of the smallest enclosing circle (Young 1988). For the Reock (Schwartzberg) test, larger (smaller) ratios indicate greater compactness. Other measures relate the activity of a district with the total activity of all basic units within the smallest enclosing circle (Ricca and Simeone 2008) or determine the ratio of the squared diameter of a district and its area (Garfinkel and Nemhauser 1970). A common global measure for the compactness of a districting plan is based on the length of the boundary between districts, i.e., the total length of the perimeter of the districts in the interior (Bozkaya et al. 2003; Lei et al. 2012)

$$cmp(\mathcal{D}) = \sum_{k=1}^p P(D_k) - P(J).$$

Short inter-district boundaries typically result in compact districts. Numerous other measures have been discussed in the literature. Unfortunately, none of them is comprehensive; some fail to detect districts that are obviously noncompact, others assign a low rating to visibly compact districts (Niemi et al. 1990; Horn et al. 1993; Williams 1995).

To use geometric measures for basic units representing points or lines, one can try to give “shape” to the districts, for example through the smallest enclosing rectangle or circle, or through the convex hull. Instead of the convex hull, one can also use χ -shapes, which are polygons enclosing the point set that can provide a better fit to the points than the convex hull (Duckham et al. 2008). However, much more common are the following, distance-based measures:

23.4.4.2 Distance-Based Measures

Distance-based measures are used predominantly in applications where people have to travel within the districts, e.g., sales- or mailmen. This confers with the motivation of compact districts in these applications: to reduce the day-to-day travel times. Moreover, in these applications basic units typically represent points or lines, making geometric measures unapplicable in the first place. The most common group of local measures is based on the sum of distances between the center of a district

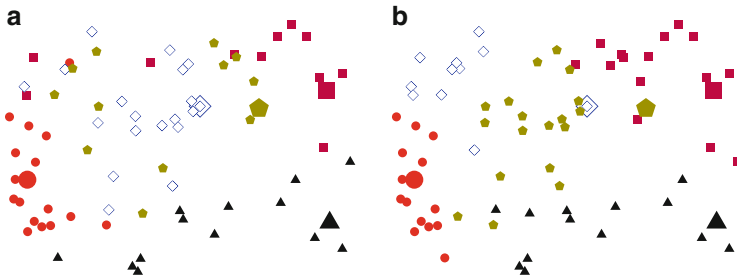


Fig. 23.3 Districting plans for two center-based compactness measures without contiguity. (a) Districts for $cmp_{ud}(\cdot)$. (b) Districts for $cmp_{wd^2}(\cdot)$

and its basic units. Variations exist in whether the distances are weighted with activity measures or not (w/u) and whether distances are squared or not (d^2/d)

$$\begin{aligned}
 cmp_{ud}(D_k) &= \sum_{j \in D_k} d_{c_k,j} & cmp_{ud^2}(D_k) &= \sum_{j \in D_k} d_{c_k,j}^2 \\
 cmp_{wd}(D_k) &= \sum_{j \in D_k} w_j d_{c_k,j} & cmp_{wd^2}(D_k) &= \sum_{j \in D_k} w_j d_{c_k,j}^2
 \end{aligned}$$

(Bard and Jarrah 2009; Bergey et al. 2003; Hess and Samuels 1971; Zoltners and Sinha 2005). The second and fourth measure are also known as the (weighted) moment of inertia (Hess et al. 1965). Although the four local compactness measures follow the same idea, the resulting districts may look considerably different as the following example shows.

Example 23.2 Consider a point set of $n = 75$ basic units that has to be partitioned into $p = 5$ districts, each having a predetermined center. The allowed relative deviation in terms of balance from the mean district size μ is 5%, and contiguity is not explicitly imposed. Figure 23.3 shows the resulting districting plans that minimize the sum of the two center-based compactness measures $cmp_{ud}(\cdot)$ and $cmp_{wd^2}(\cdot)$ over all districts. The enlarged icons represent the district centers.

Having in mind that compactness acts as a proxy for travel times, the most natural measure is $cmp_{ud}(\cdot)$. However, we observe that there is a considerable overlap in the districts for this measure, especially between the districts represented by the diamond and pentagon shaped basic units. A much better visual separation is instead obtained for the weighted squared distance, $cmp_{wd^2}(\cdot)$, even if some district centers now lie outside their actual district (again, diamonds and pentagons). A large overlap between districts typically yields less efficient routes for sales persons. To underline this observation, we determine for each district the TSP tour through all basic units, including the center. The total lengths of the TSP tours for the two districting plans are: 92.78 and 73.56. The travel distances for the weighted squared distance are 20% smaller than for $cmp_{ud}(\cdot)$. The results for $cmp_{wd}(\cdot)$ and $cmp_{ud^2}(\cdot)$ in terms of

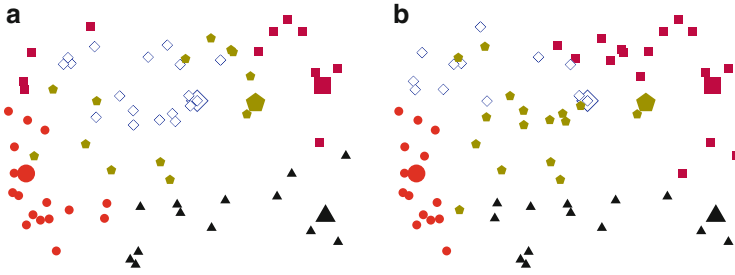


Fig. 23.4 Districting plans for two center-based compactness measures with contiguity. (a) Districts for $cmp_{ud}(\cdot)$. (b) Districts for $cmp_{wd^2}(\cdot)$

overlap and travel distances are between the other two measures, with the former being slightly better.

The situation is different if we try to enforce contiguity. Assume that an approximate neighborhood graph has been computed using the approach in Haugland et al. (2007). Using the contiguity constraints of Shirabe (2009), the resulting districting plans for $cmp_{ud}(\cdot)$ and $cmp_{wd^2}(\cdot)$ are shown in Fig. 23.4. The separation between the districts for $cmp_{ud}(\cdot)$ is clearer than before. However, even if the total length of the TSP tours reduces considerably (from 92.78 to 81.15), the districts consisting of the diamond, pentagon, and square shaped basic units are still distorted and will receive little approval from planners. (The square shaped district is connected since there exists an edge along the top of the point set.) For $cmp_{wd^2}(\cdot)$ the overlap is not much different from the previous plan, and the total travel distance even slightly decreased to 72.97. The main difference is that the centers are now all included in their districts, if only at the boundary.

This example illustrates the considerable differences between districting plans for different compactness measures and the influence of contiguity constraints. However, this is just a single example, and the observations cannot be generalized without further testing. Also, the length of a TSP tour is just an indicator for travel distances, as a sales person may not visit all customers on a single day.

The fact that squared distances produce compact but non-contiguous districts for fixed centers has been observed several times in the past (Hojati 1996; Schröder 2001). An important factor influencing the shape of districts is the spatial distribution of the district centers. If they are spread evenly, the differences between the measures in terms of district overlap will decrease, see Example 23.3. However, this uneven distribution is not unusual as sales force residences often concentrate in certain areas, e.g., larger cities, and sometimes even have the same address. Also the threshold for the allowed balance deviation has an impact on the compactness of solutions. The smaller the threshold value is, the larger the overlap between districts will get.

Instead of taking the sum, one could also take the maximum for each of the center-based measures (cf. Elizondo-Amaya et al. 2014; Ríos-Mercado and

Fernández 2009; Muyldermans et al. 2003). However, this leaves considerable freedom for assignments below the maximal distance and typically increases the overlap. A slightly different approach is based on the maximal pairwise distance and the weighted sum of pairwise distances

$$cmp_{mpw}(D_k) = \max_{i,j \in D_k, i \neq j} d_{ij} \quad cmp_{spw}(D_k) = \sum_{i,j \in D_k, i \neq j} w_i w_j d_{ij}$$

(Ríos-Mercado and Salazar-Acosta 2011 and Blais et al. 2003, respectively).

In case of measures based on the sum (maximum) of distances, the global compactness of a districting plan is then usually also computed as the sum (maximum) over all districts. But sometimes also a sum-max combination is used or a convex combination of sum and max (Muyldermans et al. 2003; Silva de Assis et al. 2014; Butsch et al. 2014).

23.4.4.3 Mathematical Modelling

The majority of districting models has compactness as an objective function to be optimized. In addition, sometimes the maximal distance between a basic unit and its district center or between two basic unit of the same district is restricted (Benzarti et al. 2013). The appeal of distance-based measures is that they easily give rise to linear or, in case of pairwise distances, quadratic expressions. Therefore, these measures are sometimes also used for polygonal basic units, even if geometric measures could have been applied (Ríos-Mercado and Fernández 2009).

23.4.5 District Center

Strictly speaking, determining district centers is in most cases not an optimization criterion in itself. However, several measures for contiguity and compactness rely on district centers. Thus, if no centers are predefined for the districts, seeking district centers is part of the optimization process. Typically, a district center is the basic unit of the district that minimizes the respective compactness measure. But also the (weighted) center of gravity can be used to determine a district center. Note however that this center usually does not coincide with a basic unit, which is problematic if distance computations are based on road networks.

23.4.6 Other Criteria

There are a few other criteria for districting problems that are included from time to time in districting models. For example, for re-districting problems the changes in

allocation from the old to the new districting plan should be kept small (Silva de Assis et al. 2014). Especially in sales territory design, customers often have a preferred sales representative by whom they want to be serviced or vice-versa, i.e., customers have banned salesmen (cf. Ríos-Mercado and López-Pérez 2013). Another criterion concerns the number of districts. Typically, p is predetermined such that, for example, the expected workload in a district neither exceeds the working time restriction of a deliverer nor renders him underutilized. If however travel times within a district account for a large portion of the total working time, then it is not always possible to fix p a priori since travel times strongly depend on the shape of districts, i.e., their compactness. Therefore, sometimes p is a design criterion (cf. Muyldermans et al. 2003).

23.5 Solution Approaches

As with most optimization problems also for districting many different solution approaches have been proposed in the literature over the years. These approaches can roughly be divided in those that utilize a mathematical programming model and those that depend merely upon heuristics. Among the former, location-allocation and set partitioning methods have been discussed. The latter mainly focus on geometric algorithms, simple construction methods, and classical meta heuristics, like Tabu Search, GRASP, and Simulated Annealing. This section will present only a rough overview and description of the most common approaches. Detailed reviews can be found in Kalcsics et al. (2005) and Ricca et al. (2013).

23.5.1 Location-Allocation Methods

The first mathematical programming approach was proposed by Hess et al. (1965) for political districting. They had the idea to model the problem as a capacitated p -median facility location problem (see also Chap. 3). Basic units correspond to customers and their activity measure to their demand. The facilities to be located are the district centers, and the capacity of the facilities is chosen in such a way that the districts obtained by solving the problem are well balanced. Candidate locations for the facilities are all basic units. For an allowed relative deviation $\alpha > 0$ of the district size from the mean district size μ , the formulation of Hess et al. (1965) is

$$\text{minimize} \quad \sum_{i,j \in J} w_j d_{ij}^2 x_{ij} \quad (23.1)$$

$$\text{subject to} \quad \sum_{i \in J} x_{ij} = 1 \quad \forall j \in J \quad (23.2)$$

$$\sum_{j \in J} w_j x_{ij} \geq (1 - \alpha) \mu y_i \quad \forall i \in J \quad (23.3)$$

$$\sum_{j \in J} w_j x_{ij} \leq (1 + \alpha) \mu y_i \quad \forall i \in J \quad (23.4)$$

$$\sum_{i \in J} y_i = p \quad (23.5)$$

$$y_i, x_{ij} \in \{0, 1\} \quad \forall i, j \in J, \quad (23.6)$$

where $x_{ij} = 1$ if basic unit j is assigned to basic unit i , 0 otherwise, and $y_i = 1$ if basic unit i is selected as district center, 0 otherwise. The objective function (23.1) maximizes the compactness of the districts using the center-based measure $cmp_{wd^2}(\cdot)$. Constraints (23.2), together with the integrality constraints on the x_{ij} -variables, model the unique and exclusive assignment criterion. Constraints (23.3) and (23.4) restrict the balance of the districts. Finally, Constraints (23.5) ensure that exactly p basic units are selected as district centers. As a result, all basic units allocated to the same basic unit i constitute a district with the basic unit as its center, i.e., there is a one-to-one correspondence between centers and districts. Note that the centers are just required to evaluate district compactness and have no meaning in itself.

Unfortunately, due to its NP-hardness, the practical use of this formulation is limited to instances with a few hundred basic units, which is rather small for typical sales districting problems. To this end, Hess et al. (1965) propose to use Cooper's location-allocation heuristic to solve the problem. In this heuristic, the simultaneous location and allocation decisions of the underlying facility location problem are decomposed into two independent phases, a location and an allocation phase, which are alternatingly performed until a satisfactory result is obtained. In the location phase, a set J_c of district centers is determined. A fairly simple and commonly used method is to solve in each district resulting from the last allocation phase a single facility location problem with the respective compactness measure as objective function (cf. Fleischmann and Paraschis 1988; George et al. 1997). To obtain an initial set of centers, one can determine new centers based on the solution of a Lagrangean subproblem (Hojati 1996). Alternatively, one can use any of the heuristics for the (uncapacitated) p -median problem or one of the heuristics mentioned below.

Once the centers have been fixed, the allocation phase determines a balanced assignment of basic units to district centers. This can be done by fixing $y_i = 1$ for all $i \in J_c$ in the above formulation. With present-day computers and MIP solvers, the resulting problem can be solved optimally even for large instances with 10,000 basic units or more within a short time. Even in the presence of contiguity constraints, several thousand basic units can be assigned in reasonable time (Ríos-Mercado and López-Pérez 2013). Alternatively, the allocation problem can be modeled as a minimum cost network flow problem allowing more flexibility for measuring and optimizing the balance and compactness of districts (George et al. 1997).

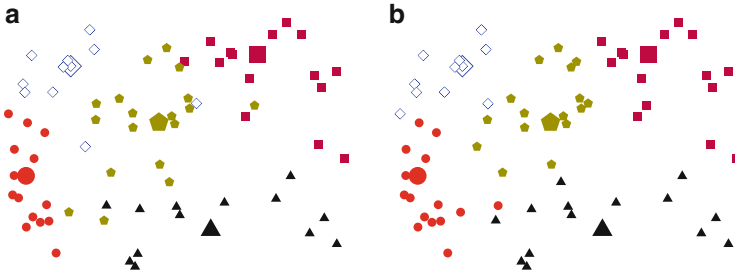


Fig. 23.5 Illustration of one iteration of the location-allocation procedure. (a) Location phase: new districts centers. (b) Allocation phase: new districts

Example 23.3 Consider again the example depicted in Fig. 23.3, but assume now that the district centers are flexible and the current ones are just a starting point. Based on the districting plan for the measure $cmp_{wd^2}(\cdot)$, the new centers that minimize $cmp_{wd^2}(\cdot)$ over each district are shown on the left-hand side in Fig. 23.5. The subsequent allocation phase yields the new districts shown on the right-hand side. The districts are visually much more compact and there is no overlap between the convex hulls of the districts.

In former times, when the exact solution of the allocation problem was unattainable for larger instances, the assignment problem was solved heuristically. Setting the tolerance α to zero and relaxing the integrality constraints on the assignment variables, i.e., $x_{ij} \in [0, 1]$, the resulting linear program is a classical transportation problem that can be solved efficiently using specialized network algorithms. However, solving the relaxed problem yields districts that are perfectly balanced but usually assign portions of basic units to more than one district, i.e., $\exists i, i' \in J_c, i \neq i', j \in J$, such that $x_{ij}, x_{i'j} > 0$. Such basic units are called splits. For an optimal basic feasible solution of the transportation problem, it is easy to prove that there are at most $p - 1$ splits (Hojati 1996). To restore the integrity of basic units, it is necessary to round for every split its fractional variables to one (one variable) or zero (the other variables). This yields disjoint districts but destroys their perfect balance. A simple split resolution rule is to assign a split to the district (center) that “owns” the largest share of the split (Hess and Samuels 1971). However, if there are just few basic units per district, this rule may produce very unbalanced districts. An optimal split allocation with a minimal maximal percentage deviation can be obtained in polynomial time by using tree partitioning methods; unfortunately, the problem of finding a split resolution with a minimal total deviation is NP-hard; see Schröder (2001) for details.

23.5.2 *Set-Partitioning Models*

As districting is essentially a partitioning problem, classical set partitioning approaches can be used to solve the problem. In a first step, balanced, contiguous, and compact candidate districts are generated in a heuristic fashion. In a second step, districts are selected from the set of candidates to optimize the overall balance of the district plan (Garfinkel and Nemhauser 1970; Mehrotra et al. 1998). Unfortunately, only small instances can be solved optimally with this approach. An advantage compared to location-allocation methods is however that almost any criterion can be applied on the generation of candidate districts.

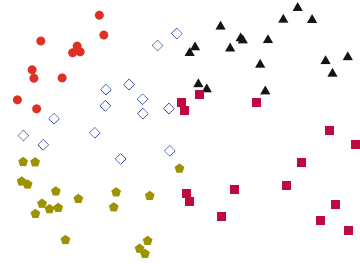
23.5.3 *Computational Geometry Methods*

A very simple but efficient solution approach for basic units representing points is the successive dichotomies strategy (Kalcsics et al. 2005). The main idea is to recursively subdivide the problem geometrically using lines into smaller and smaller subproblems until an elementary level is reached, where the problem can be solved efficiently. Hence, the basic operation is to partition a subset J' of basic units into two subsets J'_l and J'_r by drawing a line within this set of points. Given a number of line directions, for each direction the position of the line is determined in such a way that the two resulting subproblems are best balanced. For every direction, the line is evaluated by a convex combination of its balance and its compactness (evaluated through the length of inter-district boundaries), and the best line is then used to divide the problem into two subproblems. This procedure is repeated until every subset corresponds to a single district. The strategy quickly determines a well-balanced districting plan with no overlap between districts. However, as it does not explicitly account for (road) distances, the resulting districts sometimes lack compactness. Moreover, it is difficult to include neighborhood information. Instead of using lines, other geometric concepts can be used. Alternatively, the process of subdividing a point set J' can be modeled and solved as a 2-facility location problem (Salazar-Aguilar et al. 2013a).

Example 23.4 Consider again the example in Fig. 23.3 and assume that the district centers are flexible. Figure 23.6 shows the districting plan obtained with the successive dichotomies algorithm using horizontal, vertical, and diagonal lines.

Another approach is based on weighted Voronoi diagrams on networks (for a definition of weighted Voronoi diagrams see Aurenhammar et al. 2013). Assume that the neighborhood graph G is given. For center-based measures the most compact solution is obtained by assigning each basic unit to the closest center. If the distances $\{d_{c_k,j} \mid c_k \in J_c\}$ are unique for each $j \in J$, then each district will also be connected. However, the resulting districts are often far from being balanced. To overcome this drawback, the idea is to modify the distances $d_{c_k,j}$

Fig. 23.6 Districting plan with the successive dichotomies algorithm



between basic units and centers in such a way that assignments to overly large districts are “penalized” and allocations to too small districts are “stipulated”. There are basically two options to modify distances. The first adds a real-valued weight $r_k \in \mathbb{R}$ to each distance $d_{c_k,j}$ (Zoltners and Sinha 1983) and the second multiplies $d_{c_k,j}$ by a positive weight $r_k \in \mathbb{R}^+$ (Ricca et al. 2008). Hence, basic unit $j \in J$ is closer to center c_k than to center $c_l \in J_c$ if $d_{c_k,j} + w_k < d_{c_l,j} + w_l$ or $w_k d_{c_k,j} < w_l d_{c_l,j}$, respectively. Increasing (decreasing) the weight for a specific center c_k while keeping the other weights unchanged, will reduce (increase) the number of basic units assigned to c_k under the closest assignment rule and thus reduce (increase) the size of the district. To obtain balanced districts, the weights have to be updated iteratively until a satisfactory result is obtained. During the update, care has to be taken because some districts may turn out empty under additive weights or become disconnected for multiplicative weights if the weights are too uneven. For details on the update procedures see Zoltners and Sinha (1983) and Ricca et al. (2008). The partitions of the graph induced by these weights are the so-called additively and multiplicatively weighted Voronoi diagrams. Note that the approach using additive weights is in fact a Lagrangean relaxation where the balancing constraints have been relaxed.

Most districting problems are solved using discrete models. However, these problems (and a number of other logistics problems as well) can be converted into problems with continuous demand functions. Continuous demand approximations models are based on the spatial density and distribution of demand rather than on precise information on every demand point. Given continuous approximations, one can for example use Voronoi diagrams to compute or to smooth existing districts (Galvão et al. 2006), or determine perfectly balanced districts (Carlsson and Delage 2013).

23.5.4 Construction Methods

There exist several easy approaches for constructing a districting plan from scratch. One of the most popular ones is based on the multi-kernel growth methodology first introduced in Vickrey (1961). The general idea of this methodology is to select a certain number of basic units as “seed centers” and then assign to each seed

neighboring basic units in order of decreasing distance until the desired district size is reached. Variations exist with respect to the selection of seeds, whether districts grow simultaneously or sequentially around the seeds, and how to deal with enclaves of unassigned basic units which typically occur at the end of this greedy process (Bodin and Levy 1991; Williams 1995; Mehrotra et al. 1998; Bozkaya et al. 2003). The resulting districting plans are not always connected or balanced and typically serve as a starting point for a meta heuristic.

A different approach treats each basic unit initially as a single district and then merges iteratively pairs of districts until the prescribed number of districts is reached (Deckro 1977).

23.5.5 *Meta Heuristics*

There exists a wide range of meta heuristics that have been applied to districting problems: Simulated Annealing (D'Amico et al. 2002), Tabu Search (Ricca and Simeone 2008; Bozkaya et al. 2003), GRASP (Ríos-Mercado and Fernández 2009; Salazar-Aguilar et al. 2013b), and Genetic algorithms (Forman and Yue 2003; Bergey et al. 2003; Bação et al. 2005), just to name a few. A major advantage of these methods is their flexibility to include almost any practical criterion and measure for the design of districts.

23.6 Conclusions

Despite the large number of publications, it is striking that only few authors consider the districting problem independently from a practical background. Moreover, there is no consensus on which criteria are eligible and important and, on how to measure them appropriately. Thus, instead of devising yet another (variant of a) meta heuristic for a districting model with yet another measure for compactness or additional constraint, research should foremost concentrate on a common and generic framework for districting problems. And it should try to categorize the suitability of criteria and measures based on the availability of data, the geometric representation of the basic units, and the different types of applications.

Acknowledgements This work was partly supported by grant NI 521/6-1 of the German Research Foundation (DFG). This support is gratefully acknowledged.

References

- Aurenhammar F, Klein R, Lee DT (2013) Voronoi diagrams and Delaunay triangulations. World Scientific, Singapore
- Baço F, Lobo V, Painho M (2005) Applying genetic algorithms to zone design. *Soft Comput* 9:341–348
- Baker J, Clayton E, Moore L (1989) Redesign of primary response areas for county ambulance services. *Eur J Oper Res* 41:23–32
- Bard JF, Jarrah AI (2009) Large-scale constrained clustering for rationalizing pickup and delivery operations. *Transp Res B Methodol* 43:542–561
- Benzarti E, Sahin E, Dallery Y (2013) Operations management applied to home care services: analysis of the districting problem. *Decis Support Syst* 55:587–598
- Bergey P, Ragsdale C, Hoskote M (2003) A simulated annealing genetic algorithm for the electrical power districting problem. *Ann Oper Res* 121:33–55
- Blais M, Lapierre S, Laporte G (2003) Solving a home-care districting problem in an urban setting. *J Oper Res Soc* 54:1141–1147
- Bodin L, Levy L (1991) The arc partitioning problem. *Eur J Oper Res* 53:393–401
- Bozkaya B, Erkut E, Laporte G (2003) A tabu search heuristic and adaptive memory procedure for political districting. *Eur J Oper Res* 144:12–26
- Bozkaya B, Erkut E, Haight D, Laporte G (2011) Designing new electoral districts for the city of Edmonton. *Interfaces* 41:534–547
- Butsch A, Kalcsics J, Laporte G (2014) Districting for arc routing. *INFORMS J Comput* 26:809–824
- Carlsson J, Delage E (2013) Robust partitioning for stochastic multivehicle routing. *Oper Res* 61:727–744
- D'Amico S, Wang S, Batta R, Rump C (2002) A simulated annealing approach to police district design. *Comput Oper Res* 29:667–684
- Deckro R (1977) Multiple objective districting: a general heuristic approach using multiple criteria. *Oper Res Q* 28:953–961
- Dixon RJ (1968) Democratic representation: reapportionment in law and politics. Oxford University Press, New York
- Drexler A, Haase K (1999) Fast approximation methods for sales force deployment. *Manag Sci* 45:1307–1323
- Duckham M, Kulik L, Worboys M, Galton A (2008) Efficient generation of simple polygons for characterizing the shape of a set of points in the plane. *Pattern Recognit* 41:3224–3236
- Elizondo-Amaya M, Ríos-Mercado R, Díaz J (2014) A dual bounding scheme for a territory design problem. *Comput Oper Res* 44:193–205
- Ferland J, Guénette G (1990) Decision support system for a school districting problem. *Oper Res* 38:15–21
- Fleischmann B, Paraschis J (1988) Solving a large scale districting problem: a case report. *Comput Oper Res* 15:521–533
- Forman S, Yue Y (2003) Congressional districting using a TSP-based genetic algorithm. In: Proceedings of the 2003 international conference on genetic and evolutionary computation: Part II, GECCO'03, pp 2072–2083
- Galvão L, Novaes A, Souza de Cursi J, Souza J (2006) A multiplicatively-weighted Voronoi diagram approach to logistics districting. *Comput Oper Res* 33:93–114
- Garfinkel R, Nemhauser G (1970) Optimal political districting by implicit enumeration techniques. *Manag Sci* 16:495–508
- George J, Lamar B, Wallace C (1997) Political district determination using large-scale network optimization. *Socio Econ Plan Sci* 31:11–28
- Glaze T, Weinberg C (1979) A sales territory alignment program and account planning system. In: Sales management: new developments from behavioral and decision model research. Marketing Science Institute, Cambridge, pp 325–343

- Gross J, Yellen J (2003) Handbook of graph theory. CRC Press, Boca Raton
- Hanafi S, Fréville A, Vaca P (1999) Municipal solid waste collection: an effective data structure for solving the sectorization problem with local search methods. *INFOR* 37:236–254
- Handley L, Grofmann B (eds) (2008) Redistricting in comparative perspective. Oxford University Press, New York
- Haugland D, Ho S, Laporte G (2007) Designing delivery districts for the vehicle routing problem with stochastic demands. *Eur J Oper Res* 180:997–1010
- Hess S, Samuels S (1971) Experiences with a sales districting model: criteria and implementation. *Manag Sci* 18:41–54
- Hess S, Weaver J, Siegfeldt H, Whelan J, Zitlau P (1965) Nonpartisan political redistricting by computer. *Oper Res* 13:998–1008
- Hojati M (1996) Optimal political districting. *Comput Oper Res* 23:1147–1161
- Horn D, Hampton C, Vandenberg A (1993) Practical application of district compactness. *Polit Geogr* 12:103–120
- Howick R, Pidd M (1990) Sales force deployment models. *Eur J Oper Res* 48:295–310
- Jarrah A, Bard J (2012) Large-scale pickup and delivery work area design. *Comput Oper Res* 39:3102–3118
- Kalcsics J, Nickel S, Schröder M (2005) Towards a unified territorial design approach – applications, algorithms and GIS integration. *TOP* 13:1–74
- Lei H, Laporte G, Guo B (2012) Districting for routing with stochastic customers. *EURO J Transp Logist* 1:67–85
- Lewyn M (1993) How to limit gerrymandering. *Florida Law Rev* 45:403–486
- Lin HY, Kao JJ (2008) Subregion districting analysis for municipal solid waste collection privatization. *J Air Waste Manag Assoc* 58:104–111
- Lodish L (1975) Sales territory alignment to maximize profit. *J Market Res* 12:30–36
- López-Pérez J, Ríos-Mercado R (2013) Embotelladoras ARCA uses operations research to improve territory design plans. *Interfaces* 43:209–220
- Mehrotra A, Johnson E, Nemhauser G (1998) An optimization based heuristic for political districting. *Manag Sci* 44:1100–1114
- Minciardi R, Puliafito PP, Zoppi R (1981) A districting procedure for social organizations. *Eur J Oper Res* 8:47–57
- Muyldermans L, Cattrysse D, Van Oudheusden D, Lotan T (2002) Districting for salt spreading operations. *Eur J Oper Res* 139:521–532
- Muyldermans L, Cattrysse D, Van Oudheusden D (2003) District design for arc-routing applications. *J Oper Res Soc* 54:1209–1221
- Nagel S (1965) Simplified bipartisan computer redistricting. *Stanford Law Rev* 17:863–899
- Niemi R, Grofman B, Carlucci C, Hofeller T (1990) Measuring compactness and the role of a compactness standard in a test for partisan and racial gerrymandering. *J Polit* 52:1155–1181
- Parker F (1990) Black votes count. The University of North Carolina Press, Chapel Hill
- Puppe C, Tasnádi A (2008) A computational approach to unbiased districting. *Math Comput Model* 48:1455–1460
- Ricca F, Simeone B (2008) Local search algorithms for political districting. *Eur J Oper Res* 189:1409–1426
- Ricca F, Scozzari A, Simeone B (2008) Weighted Voronoi region algorithms for political districting. *Math Comput Model* 48:1468–1477
- Ricca F, Scozzari A, Simeone B (2013) Political districting: from classical models to recent approaches. *Ann Oper Res* 204:271–299
- Ríos-Mercado R, Fernández E (2009) A reactive GRASP for a commercial territory design problem with multiple balancing requirements. *Comput Oper Res* 36:755–776
- Ríos-Mercado R, López-Pérez J (2013) Commercial territory design planning with realignment and disjoint assignment requirements. *Omega* 41:525–535
- Ríos-Mercado R, Salazar-Acosta J (2011) A GRASP with strategic oscillation for a commercial territory design problem with a routing budget constraint. *Advances in Soft Computing Lecture notes in computer science, 10th Mexican International Conference on Artificial Intelligence,*

- MICAI 2011, Puebla, Mexico, Proceedings, Part II, vol 7095, pp 307–318, Springer Berlin Heidelberg
- Salazar-Aguilar M, González-Velarde J, Ríos-Mercado R (2013a) A divide-and-conquer approach for a commercial territory design problem. *Computación y Sistemas* 16:309–320
- Salazar-Aguilar M, Ríos-Mercado R, González-Velarde J (2013b) GRASP strategies for a bi-objective commercial territory design problem. *J Heuristics* 19:179–200
- Schoepfle O, Church R (1991) A new network representation of a “classic” school districting problem. *Socio Econ Plan Sci* 25:189–197
- Schröder M (2001) Gebiete optimal aufteilen. Ph.D. thesis, Universität Karlsruhe. <http://www.ubka.uni-karlsruhe.de/eva>
- Shirabe T (2009) Districting modeling with exact contiguity constraints. *Environ Plan B Plan Des* 36:1053–1066
- Silva de Assis L, Morelato Franca P, Luiz Usberti F (2014) A redistricting problem applied to meter reading in power distribution networks. *Comput Oper Res* 41:65–75
- Simchi-Levi D, Kaminsky P, Simchi-Levi E (2003) *Designing & managing the supply chain: concepts, strategies & case studies*, 2nd edn. McGraw-Hill/Irwin, New York
- Skiera B, Albers S (1994) COSTA: Ein Entscheidungs-Unterstützungs-System zur deckungsbeitragsmaximalen Einteilung von Verkaufsgebieten. *Z Betriebswirt* 64:1261–1283
- Tavares-Pereira F, Figueira J, Mousseau V, Roy B (2007) Multiple criteria districting problems: the public transportation network pricing system of the Paris region. *Ann Oper Res* 154:69–92
- Vickrey W (1961) On the prevention of gerrymandering. *Polit Sci Q* 76:105–110
- Webster G (2013) Reflections on current criteria to evaluate redistricting plans. *Polit Geogr* 32:3–14
- Williams JC Jr (1995) Political redistricting: a review. *Pap Reg Sci* 74:13–40
- Xu C, Yum TSP (2010) Patrol districting and routing with security level functions. In: *Proceedings of the IEEE international conference on systems, man and cybernetics*, pp 3555–3562
- Young H (1988) Measuring the compactness of legislative districts. *Legis Stud Q* 13:105–115
- Zhong H, Hall R, Dessouky M (2007) Territory planning and vehicle dispatching with driver learning. *Transp Sci* 41:74–89
- Zhou G, Min H, Gen M (2002) The balanced allocation of customers to multiple distribution centers in a supply chain network: a genetic algorithm approach. *Comput Ind Eng* 43:251–261
- Zoltners AA, Sinha P (1983) Sales territory alignment: a review and model. *Manag Sci* 29:1237–1256
- Zoltners A, Sinha P (2005) Sales territory design: thirty years of modeling and implementation. *Market Sci* 24:313–331