OR Spectrum
https://doi.org/10.1007/s00291-024-00745-3

ORIGINAL ARTICLE

Police service district planning

Tobias Vlćek[1] · **Knut Haase**[1] · **Malte Fliedner**[2] · **Tobias Cors**[2]

Received: 1 June 2023 / Accepted: 8 January 2024

Abstract

We propose a new framework to address the territory design problem of emergency services in collaboration with two police authorities in Europe. Our framework serves as a strategic decision support system to assess different districting layouts, department locations, staffing decisions and dispatching strategies. First, we introduce a novel modification of the *p*-median problem with a combined approach to the contiguity and compactness of district layouts solvable by a commercial solver. Second, we utilize a new discrete event simulation that accounts for the variability of spatial and temporal incident patterns and driving times to evaluate the district layouts according to several criteria based upon up to 1.8 million historical incidents. Our simulation results demonstrate that our proposed district layouts can lead to a reduction of the response time by up to 14.52% while also lowering the dispatch time, the overall driving time, and the number of unanswered calls for service. Additionally, we examine the computational complexity of optimally locating district centers and analyze the more restricted problem of optimally reassigning districts to fixed district centers.

Keywords Decision support systems · Territory design · Emergency services · Simulation

✉ Tobias Vlćek
tobias.vlcek@uni-hamburg.de

Knut Haase
knut.haase@uni-hamburg.de

Malte Fliedner
malte.fliedner@uni-hamburg.de

Tobias Cors
tobias.cors@uni-hamburg.de

[1] Institute of Logistics, Transport and Production, Universität Hamburg, Moorweidenstraße 18, 20148 Hamburg, Germany

[2] Institute of Operations Management, Universität Hamburg, Moorweidenstraße 18, 20148 Hamburg, Germany

Published online: 10 February 2024

1 Introduction

Police services have to fulfill the needs of three interest groups: citizens, service personnel and administrators (Bodily 1978). From a citizen's perspective, the primary aim is a fast and successful response to calls for service to reduce loss of life, loss of property and injuries (Swersey 1994). From the perspective of the service personnel, the principal objectives are equal risks and workloads (Camacho-Collados and Liberatore 2015). The administrators have to reconcile all objectives. Hence, they aim for an efficient and effective service operation, where 'efficient' refers to low operational expenses and 'effective' describes a high quality of service (Shen et al. 1995). However, there is a financial trade-off between an efficient and an effective response (de Souza et al. 2015). Furthermore, administrators often stand under political pressure and resource constraints (Curtin et al. 2005). The improved utilization of information aids police service administrators to meet their objectives while addressing all needs. Districting optimization is a part of this data-driven improvement process. For instance, many police district layouts date back several decades. In the past, administrators designed them manually to run along major streets, highways and regional boundaries (Bruce 2009). An efficient and effective response was usually not a criterion during the procedure (Mitchell 1972). But in recent decades, these layouts attracted the attention of researchers. Through the use of operations research (OR) models, incident records and geographical information, administrators can improve the response of police services without an increase in operational expenses. Besides the optimization of district layouts and patrol beats, OR models help administrators in making strategic decisions regarding the location of new departments or the closure of certain departments because of budgetary constraints (Liberatore et al. 2020; Cheung et al. 2015). Despite these benefits, the adoption of OR techniques to improve the district layouts of police services has been slow (Green and Kolesar 2004).

We contribute to the existing research by allocating police departments and district layouts at the same time. Contrary to previous modeling approaches, we use a new set of constraints to simultaneously address the contiguity and compactness of districts. We further demonstrate in our study that the increased compactness of our novel constraints results in a better distribution of police departments. Furthermore, we examine the complexity of allocating district centers and use a new simulation to evaluate additional factors influencing the police response.

2 Background

For an efficient and effective distribution of resources, police jurisdictions are divided into precincts or command districts with separate departments. These are further divided into patrol beats (D'Amico et al. 2002). In our case studies for police authorities in Germany and Belgium, a part of the police force patrols the streets, and another part is stationed at the departments to respond to calls for service (CFS) of different priorities. The highest priority is assigned to severe incidents, such as murders in progress or hostage-situations. As these incidents are critical, more personnel has to be dispatched because staff shortages put citizens' and officers' lives at risk. The lowest priority is reserved for non-critical incidents, such as disturbances of the peace or minor traffic accidents. Dispatchers assign all CFS to vehicles from the corresponding districts and patrol areas, as the officers are familiar with the area and are thus better prepared to respond appropriately (Bodily 1978). To cope with high demands, dispatchers can assign vehicles from nearby districts or beats. However, this can lead to a domino effect, as transferring vehicles from other districts or beats reduces coverage in those locations, making them vulnerable when they need assistance themselves (Mayer 2009). If a dispatcher cannot deploy a response unit, the incident is queued and the dispatch delayed. Depending on the strategic aim, dispatchers can skip CFS with low priorities if the system is under a heavy load to aid high-priority CFS first (de Souza et al. 2015). The worst-case scenario is an overloaded system with long dispatching delays due to staff shortages. In these systems, preventive patrol is hardly possible (Miller and Knoppers 1972) and dispatchers constantly draw on patrol resources when more help is needed.

A central criterion to measure the effectiveness of police services is the response time (Dunnett et al. 2019), whose reduction is the primary goal of our case studies. It measures the time between a call for aid and the arrival at the incident location. It comprises the call length, the dispatch delay and the driving time to the incident location (Kern 1989). A low response time has several benefits: it increases the likelihood of helping citizens, and it improves confidence in the service (Bodily 1978). Other criteria relevant to the police authorities of our case studies include the time spent at the department, the exchange rate between districts and the overall driving time of the vehicles.

3 Related research

Our problem is part of the field of territory design problems (TDP), a subfield of discrete optimization (Ríos-Mercado 2020). It involves the aggregation of small geographic areas, called basic areas (BAs), into geographic clusters, called districts, so that these are acceptable according to predefined planning criteria (Zoltners and Sinha 1983). In recent decades, researchers have made many contributions to the TDP. They have developed these approaches to focus on issues in various areas, such as political districting, service territory design, reserve design and emergency service districting. Almost all models were developed with a case study to account for specific questions. To address this issue, (Kalcsics et al. 2005) summarized and unified different approaches to the TDP to establish a mutual base in the research community.

Three important criteria to assess district layouts are contiguity, balance, and compactness. Contiguous districts allow the movement between BAs without leaving the district (Kalcsics and Ríos-Mercado 2019). Balance addresses the difference in activity measures between districts. The choice of the activity measure depends on the context of the model: it can be profit, workload, travel time, etc. or a combination of several measures (Kalcsics et al. 2005). Compactness has no univocal definition; a district is commonly declared compact if it is 'somehow round-shaped and undistorted' (Kalcsics et al. 2005). Hence, this is the definition to which we will adhere. Compactness is usually enforced through distance-based measures in the objective function, which can lead to a trade-off between different objectives (Hess et al. 1965; Mehrotra et al. 1998; Shirabe 2009; Salazar-Aguilar et al. 2011; Önal et al. 2016; Wang et al. 2018; Kalcsics and Ríos-Mercado 2019). Various methods to enforce and evaluate compactness are illustrated by Li et al. (2013). Another significant aspect of TDPs is the computation time. Clearly, problems with smaller BAs lead to more accurate results. However, there is a trade-off between a granular solution and a workable calculation time (Mitchell 1972). Specifically, the general graph-partitioning problem most models can be reduced to has been proven to be NP-hard (Garey and Johnson 1979; Reese 2006). Thus, many approaches rely on heuristics to derive a solution. Hess et al. (1965) applied the first model to solve the TDP. They proposed a model based on the warehouse-location problem to design compact and balanced political districts. They developed a location-allocation heuristic to solve it, as the problem was too complex to be solved with commercial solvers. A thorough study of further literature concerning different TDPs and their criteria can be found in the review of Kalcsics and Ríos-Mercado (2019).

One of the subproblems in territory design involves emergency services. Here, the problem arises in the context of emergency medical services, fire stations and police services. Research on emergency medical services usually involves the operational issue of allocating and reallocating ambulances based on covering problems (Mayorga et al. 2013; Regis-Hernández et al 2018). The research on fire stations is mainly centered on the allocation of the fire stations (Badri et al. 1998; Yang et al. 2004; van den Berg et al. 2017). The police districting problem (PDP) specifically addresses the objectives and constraints of police services and aims for an optimal partition of a territory under the jurisdiction of a police department. The optimization models employed to solve the PDP are mainly modifications of the warehouse-location problem or the maximum-coverage problem (Liberatore et al. 2020). The first application of the PDP was developed by Mitchell (1972). He formulated a mixed-integer problem to plan the patrol beats for police vehicles according to several objectives, including low driving times and equal workloads. To solve the problem, he had to resort to a heuristic. Different important criteria for patrol sectors were proposed by Liberatore and Camacho-Collados (2016), including risk and mutual support. Liberatore et al. (2020) wrote an excellent summary of further literature regarding the PDP. Our problem addresses issues of the PDP, but differs in the aspect that most PDP contributions search for the optimal layout of patrol beats. We aim to find the optimal department locations and their contiguous and compact district borders, a problem that is the main objective of many TDPs.

Apart from the optimization, our framework includes a police service operation simulation to evaluate the resulting districts. The major contributions in this area date back many decades. A research group developed them to optimize the efficiency and effectiveness of police services in the United States. Larson (1974) developed the hypercube queuing model (HQM) to analyze vehicle location and district design problems in multi-server queuing systems. Kolesar and Walker (1975) developed a program to simulate the activities of police patrol cars to allocate capacities. The same year, Chaiken and Dormont (1978) proposed the Patrol Car Allocation Model (PCAM), a less accurate but less resource-intense program. They based it on approximate equations and dispatch queues with the aim to improve the resource allocation as well. Afterward, Kern (1989) developed a simulation allowing for complex dispatching rules and inter-sector exchanges. Simultaneously, Green and Kolesar (1989) helped to design a revised version of the PCAM with a multiple-car dispatch queueing model. Later, several researchers used the PCAM in subsequent contributions to the PDP to provide feedback during heuristic approaches to the design of patrol beats (D'Amico et al. 2002; Zhang and Brown 2013, 2014). Dunnett et al. (2019) developed a framework to help police administrators to dispatch the optimal number of officers for incident responses and validated it with a discrete event simulation. Recently, Zhu et al. (2021) proposed an optimization framework based on the HQM by assigning 81 fixed patrol beats to 6 precincts, aiming to balance their workload while improving other criteria.

Despite the prior research on the TDP and the PDP, we believe that our contribution to the existing literature is of value. First, we contribute by introducing a novel set of constraints, that can address contiguity and geometric compactness simultaneously on any districting problem with a hexagonal grid. With these constraints, we can compute contiguous and compact district layouts with a commercial solver on instances with up to 2040 BAs within hours. Second, we prove the NP-completeness of the contiguous p-median problem with fixed centers, a common problem in the field of TDPs. As far as we are aware, the proof has not existed before. Third, we expand upon previous emergency service simulation models to support districts, inter-district exchanges, follow-ups after incidents, resource planning, additional dispatching rules and a flexible approach to priority handling, to evaluate our resulting district layouts.

4 Framework

4.1 Core problem statement and complexity analysis

Before developing a mathematical model, we formulate the two core problems solved in this work and analyze their computational complexity. The foremost goal of the optimization is to minimize response times over all incidents coming from CFS. As previously stated, response time is composed of the call length, the dispatch delay and the driving time to the incident location. We assume that the average call length is fixed and depends on the police line operator. The dispatch delay is mainly a consequence of scarce capacities and not directly affected by district layouts. However, by minimizing the driving time, we reduce capacity utilization of unproductive times and thereby indirectly lower the dispatch delay in addition to its direct effect on response times. Subsequently, we refer to the response time as its variable components: the sum of the dispatch delay and the driving time to an incident. We further adopt the assumption of Curtin et al. (2010), that all vehicles are dispatched from the department responsible for the district. This is reasonable, as in our two case studies, incidents should be covered by vehicles dispatched from departments. Patrol cars have the task to mainly handle incidents discovered on patrol and are not considered in our model. In addition, vehicles have to drive to the department regularly after incidents to transport suspects and complete paperwork, or at the end of the shift. We also consider that CFS are not evenly distributed among districts and that CFS might require multiple vehicles at once.

Let e_{pj} denote the average number of CFS with priority $p \in \mathcal{P}$ that have taken place in BA $j \in \mathcal{J}$ for some fixed time horizon in the past (e.g., 1 year) and c_{pj} denote the average number of requested cars per CFS of priority p over the same period in BA j, then $e_{pj} \times c_{pj}$ represents the demand intensity for priority p of a BA j. In addition, let d_{ij} denote the average expected driving time between all BAs $i \in \mathcal{I}$ which could serve as department location and all BAs j for a single vehicle and CFS. Finally, let w_p represent weights for the different incident priorities p to underline the importance of fast responses to severe incidents, as proposed by Mitchell (1972). Then, we can determine the total weighted driving time as:

$$t_{ij} = 2d_{ij} \times \sum_{p \in \mathcal{P}} e_{pj} \times c_{pj} \times w_p \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J} \tag{1}$$

i.e., as the sum of all weighted driving times d_{ij} from the potential department locations i to the incidents in BA j and back with $\mathcal{J}$ referring to the set of all BAs and $\mathcal{I} \subseteq \mathcal{J}$ denoting the set potential department locations (district centers). Note that this calculation includes all CFS and requested vehicles in the time horizon, including the CFS that were not fulfilled. Thus, neither the available staff nor the response time in the past exert an influence. If we further assume that all BAs can be represented in a graph of a given structure, we can define the core problem which needs to be solved to determine department locations and district layouts as a contiguous p-median problem.

Contiguous p-median problem Let $G = (\mathcal{J}, \mathcal{E})$ be an undirected graph and $\mathcal{I} \subseteq \mathcal{J}$ the set of potential district centers. Further, let $t : \mathcal{I} \times \mathcal{J} \rightarrow \mathbb{N}$ denote the total driving times between potential district centers and BAs. Select a set of district centers $\mathcal{I}^s \subseteq \mathcal{I}$ with cardinality P and find a partition of the nodes of G into P connected subgraphs U_i, with $i \in U_i \; \forall i \in \mathcal{I}^s$, $U_i \cap U_{i'} = \emptyset \; \forall i, i' \in \mathcal{I}^s, i \neq i'$, such that $\sum_{i \in \mathcal{I}^s} \sum_{j \in U_i} t_{ij}$ is minimized.

In general, p-median problems have been shown to be *NP*-hard in the strong sense even in planar graphs of maximum vertex degree 3 and if all nodes of the graph are adjacent to the p-medians in the optimal solution (see Kariv and Hakimi 1979). As a consequence, imposing standard contiguity or compactness constraints will typically not improve worst-case performance of exact algorithms and the problem version of the contiguous p-median problem studied in this work can readily be shown to be *NP*-hard based on the same reduction as in Kariv and Hakimi (1979). Nevertheless, the typical lifespan of a police department is long enough that it makes sense to adapt district layouts occasionally to reflect structural changes. This naturally leads to a contiguous districting problem for a given set of fixed departments (district centers), which can be formulated as follows:

Contiguous p-median problem with fixed centers (CPMFC) Let $G = (\mathcal{J}, \mathcal{E})$ be an undirected graph and $\mathcal{I}^S \subseteq \mathcal{J}$ be a set of given district centers with cardinality P. Further, let $t : \mathcal{I}^S \times \mathcal{J} \rightarrow \mathbb{N}$ denote the total driving times between district centers and BAs. Find a partition of the nodes of G into P connected subgraphs U_i, with $i \in U_i \; \forall i \in \mathcal{I}^S, U_i \cap U_{i'} = \emptyset \; \forall i, i' \in \mathcal{I}^S i \neq i'$, such that $\sum_{i \in \mathcal{I}^S} \sum_{j \in U_i} t_{ij}$ is minimized.

Many versions of such districting or graph tree partitioning problems have been shown to be *NP*-hard. For instance, Altman (1997) shows that contiguous districting is *NP*-hard in the strong sense if population equality constraints are enforced. Cordone and Maffioli (2004) show, among other results, that contiguous districting with inclusions constraints is *NP*-hard on grid graphs for min-sum and min-max objectives over edge weights. In this work, we will prove a complementary result that shows that contiguous districting is *NP*-hard in the strong sense for *p*-median objectives even without inclusion constraints and even for typical objective values that can be derived from regular road networks as long as district intensities are considered and driving times are subject to traffic influences. To the best knowledge of the authors, this has not yet been demonstrated for districting problems and given that the result concerns a rather basic districting problem, it might have a wide range of applicability. The following theorem for the decision version of CPMFC (CPMFC-d) is proven in the appendix.

Theorem 1 *CPMFC-d is NP-complete in the strong sense, even if* $P = 2$ *and G is a grid graph.*

Proof of Theorem 1 See appendix. □

The above result can be readily extended to cover standard compactness constraints, so that both core problems solved in this work are *NP*-hard in the strong sense. Still, that doesn't necessarily mean that practice relevant problem sizes cannot be solved in a reasonable time. To the contrary, we will show in the following that solution performance is more than adequate for real-world problem sizes, as long as contiguity and compactness constraints are modeled appropriately.

4.2 Mathematical model

The first task for the mathematical model is to transform the considered area of police operations into a well-defined graph of BAs. So far, territory design problems often make use of squares, irregular-shaped polygons or streets as BAs. Given the complexity result of the previous section, the exact graph structure will not impact worst-case algorithmic performance (given that even grid graphs are hard), we base our choices on other considerations and divide the underlying area into regular, same-sized hexagons. Unlike with squares or irregular polygons, the distances to all neighboring centroids in a hexagonal grid are identical. This reduces the sampling bias from edge effects (Wang and Kwan 2018). Moreover, it gives us a more accurate representation of the driving time d_{ij} between the BAs, as we adhere to the common approach to use the centroid of each BA as a place of departure and arrival (Curtin et al. 2005; Cheung et al. 2015; Liberatore and Camacho-Collados 2016). In addition, we will utilize the special properties of hexagons to enforce compactness.

We make use of the following sets and parameters in the model:

$\mathcal{J}$: set of BAs, indexed by j and v,
$\mathcal{I}$: set of potential district centers ($\mathcal{I} \subseteq \mathcal{J}$), indexed by i,
$\mathcal{I}^s$: set of current departments ($\mathcal{I}^s \subseteq \mathcal{I}$),
$\mathcal{J}_i$: sets of BAs assignable to $i(\mathcal{J}_i \subseteq \mathcal{J})$,
$\mathcal{A}_j$: sets of BAs adjacent to BA $j(\mathcal{A}_j \subseteq \mathcal{J},)$
p : the number of district centers (departments),
$\bar{d}$: maximum feasible driving time between center i and BA j,
$\bar{d}^q$: maximum feasible driving time for a support center of district center i,
q : minimum number of other district centers around district center i,
δ : maximum number of reallocated district centers,
a_{ij} : Euclidean distance between the BA i and BA j and
t_{ij} : total weighted driving time between BA i and BA j.

We define the following variable:

$$X_{ij} = 1, \text{ if BA } j \text{ is assigned to district center (department) BA } i \text{ (otherwise } X_{ij} = 0).$$

The mathematical model can thus be formulated as follows:

$$\text{minimize} \quad \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i} t_{ij} \times X_{ij} \tag{2}$$

subject to:

$$\sum_{i \in \mathcal{I} | j \in \mathcal{J}_i} X_{ij} = 1 \quad \forall j \in \mathcal{J} \tag{3}$$

$$\sum_{i \in \mathcal{I}} X_{ii} = p \tag{4}$$

$$X_{ij} \leq X_{ii} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i \tag{5}$$

$$X_{ij} \in \{0, 1\} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i \tag{6}$$

In our objective function (2), we minimize driving time weighted by the requested number of cars in each BA. (3) guarantees that each BA j is allocated to exactly one district center. (4) fixes the number of district centers. (5) ensures that BAs can only be assigned to realized district centers. (6) is the variable domain. Moreover, we enforce a maximal driving distance $\bar{d}$ to mitigate a system overload through long driving times with the sets $\mathcal{J}_i$ where $j \in \{\mathcal{J}_i | a_{ij} < \bar{d}\}$. Note, that we also tested the p-center problem and the maximal covering location problem for our case studies, but found that the p-median solutions yielded better results in our simulation.

To improve the possibility of inter-district exchanges and prevent isolated departments, we introduce the additional constraints (7) and (8). They ensure that at least q other district centers k are within the radius $\bar{d}^q$ of a district center i to provide support units, as proposed by Liberatore and Camacho-Collados (2016).

$$\sum_{\substack{k \in \mathcal{I} | d_{ik} < \bar{d}^q \\ i \neq k}} X_{kk} \geq q \times X_{ii} \quad \forall i \in \mathcal{I} \tag{7}$$

The strength of (7) depends on the administrators' preferences and the dispatching rules. Furthermore, we employ constraint (8) to model districts with (partially) fixed departments. It secures that only up to δ departments are moved to new locations.

$$\sum_{i \in \mathcal{I}^s} X_{ii} \geq p - \delta \tag{8}$$

4.3 Contiguity and compactness

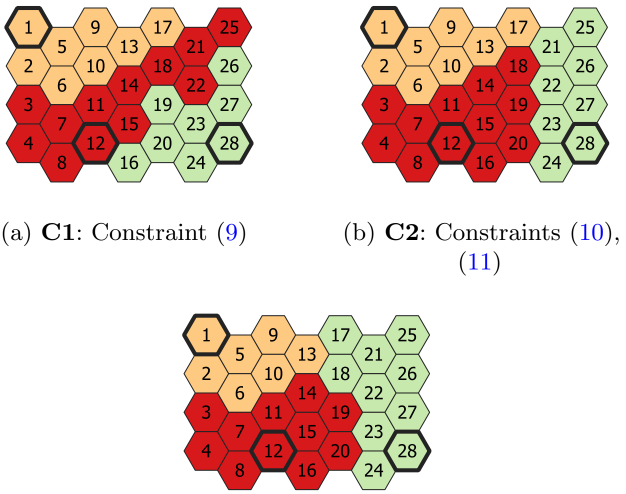
So far, we do not restrict the districts to be contiguous. Therefore, we introduce (9), which was first proposed by Zoltners and Sinha (1983).

$$X_{ij} \leq \sum_{v \in \mathcal{N}_{ij}} X_{iv} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i \setminus \mathcal{A}_i : i \neq j \tag{9}$$

Their set $\mathcal{N}_{ij}$ comprised the preceding adjacent BAs of j on any branch of a hierarchical adjacency tree for center i. The idea behind the restriction of Zoltners and Sinha (1983) is that each BA j allocated to a district center i has at least one adjacent BA v with a shorter distance to i allocated to the same center i. This ensures contiguity, as there is always a connected path from BA j back to the center i. The BAs in the research of Zoltners and Sinha were polygons of different sizes and shapes. Because our BAs are same-sized hexagons, we can simplify this restriction and use the Euclidean distance a_{ij} between the centroids of the hexagons to restrict the adjacency through the adjusted set

$$\mathcal{N}_{ij} = \{v \in \mathcal{A}_j | a_{iv} < a_{ij}\} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i.$$

We chose the restriction (9) because contiguity constraints are responsible for a considerable part of the computational effort in districting problems (Haase and Müller 2014). Wang et al. (2018) have demonstrated that a minimization of distances in the objective function together with restriction (9) contributes to a lower computation time compared to other contiguity formulations.



(a) **C1**: Constraint (9)

(b) **C2**: Constraints (10), (11)

(c) **C3**: Constraints (10), (12), (13)

Fig. 1 Contiguity constraint comparison

Moreover, we can use a variation of the approach in (9), abbreviated C1, to ensure contiguity while improving the geometric compactness of the district layouts. We do so by modifying the variable domain of (9) as displayed in (10) and we include the new constraint (11).

$$X_{ij} \leq \sum_{v \in \mathcal{N}_{ij}} X_{iv} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i \setminus \mathcal{A}_i : |\mathcal{N}_{ij}| \leq 1 \wedge i \neq j \tag{10}$$

$$2X_{ij} \leq \sum_{v \in \mathcal{N}_{ij}} X_{iv} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i \setminus \mathcal{A}_i : |\mathcal{N}_{ij}| > 1 \wedge i \neq j \tag{11}$$

The combination of both constraints, subsequently called C2, tightens the permissible solution range of the problem and excludes certain contiguous layouts possible with (9) as illustrated in the example in Fig. 1a, b.

A hexagon has six adjacent hexagons, except at the borders of a map. Accordingly, the number of adjacent BAs v closer to the district center i, $|\mathcal{N}_{ij}|$, can only be as high as three. If a BA j has less than two adjacent BAs v closer to the district center i, we actually implement the modified restriction (9) through (10). If a BA j possesses two or three adjacent BAs $v \in \mathcal{N}_{ij}$ nearer to i, we apply (11). This secures that at least two of the BAs $v \in \mathcal{N}_{ij}$ have to be allocated to i, before j can be assigned to i. While this enforces contiguity, it can likewise enhance the compactness of a district layout.

Figure 1 demonstrates three variations (a), (b) and (c) of the same area with three districts and a district center i in BA 1, 12 and 28. Figure 1a illustrates a viable contiguous solution while using constraint (9). In contrast, this layout is not allowed if we employ the restrictions (10) and (11), as BA 18, BA 22 and BA 25 violate constraint (11). They have solely one adjacent BA v closer to the district center i in BA 8. If the objective function minimizes the Euclidean distance,

$$\text{minimize} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} a_{ij} \times X_{ij},$$

such a result is unlikely, as the objective function itself aims for compact territories. However, in our model we minimize the weighted driving time t_{ij} in each BA j, which can contribute to distorted district shapes. Even stronger distortions can arise if an optimization problem is subject to additional constraints. Figure 1b presents a potential solution while using constraint (10) and (11). These districts are superior for patrols and police operations, because of their greater compactness and smaller diameter. A resulting benefit is a shorter average driving time, in the case that patrol cars serve as additional backup (Liberatore and Camacho-Collados 2016).

We can formulate a third set of contiguity constraints, called C3, to enhance the compactness even further. First, we establish the set

$$\mathcal{M}_{ij} = \{v \in \mathcal{A}_j | \exists v' \in \mathcal{A}_j : a_{iv} < a_{iv'}\} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}_i.$$

It contains all adjacent BAs $v \in \mathcal{A}_j$ except for the BA $v' \in \mathcal{A}_j$ that is the farthest away from i. Second, we include the constraint (10) and modify the variable domain of constraint (11) as displayed in constraint (12).

$$2X_{ij} \leq \sum_{v \in \mathcal{N}_{ij}} X_{iv} \quad \forall i \in \mathcal{I}, \quad \forall j \in \mathcal{J}_i \setminus \mathcal{A}_i : |\mathcal{N}_{ij}| > 1 \wedge |\mathcal{M}_{ij}| < 5 \wedge i \neq j \tag{12}$$

Third, we introduce the constraint (13), which includes all combinations of i and j excluded in (12) and (10).

$$3X_{ij} \leq \sum_{v \in \mathcal{M}_{ij}} X_{iv} \quad \forall i \in \mathcal{I}, \quad \forall j \in \mathcal{J}_i \setminus \mathcal{A}_i : |\mathcal{M}_{ij}| = 5 \wedge i \neq j \tag{13}$$

Now, our third set of constraints aims to ensure that each BA j allocated to i has at least three adjacent BAs v allocated to the same districts, if possible. (10) and (12) are applied if a BA j with the district center i has fewer than 5 elements in $\mathcal{M}_{ij}$. Here, the restrictions ensure the same contiguity and compactness as (10) and (11). If a combination of i and j in $\mathcal{M}_{ij}$ has five BAs v, we apply restriction (13). This secures that at least three BAs $v \in \mathcal{M}_{ij}$ are allocated to the same district center as j. While improving the compactness, the constraint still ensures that the resulting districts are contiguous, as the distance a_{iv} of one of those allocated BAs $v \in \mathcal{M}_{ij}$ is always lower than a_{ij}. An illustration can be seen in Fig. 1c. We intentionally excluded additional constraints to the adjacent BAs $v \in \mathcal{A}_i$ of a district center i, as the current formulation still allows small districts. Nonetheless, it is important to note that the possible district layouts resulting from C1, but especially from C2 and C3, are a subset of the set of all possible contiguous district layouts. Thus, we could find superior objective values with less restricted contiguity constraints at the cost of higher distortions to the compactness of the layouts (Mehrotra et al. 1998).

4.4 Algorithmic experiments

In the following, we will investigate the algorithmic solvability of the presented problem.

To evaluate it, we implemented our model without constraints (7) and (8) on a computer with an 8-Core AMD Ryzen 5700 G CPU and 32 GB of RAM. For our benchmark, we created four scenarios based on the same rectangular section of an urban area: one with 510 BAs, one with 1008 BAs, one with 1508 BAs and one with 2040 BAs. First, we replaced the total weighted number of requested cars in each BA by the number of street intersections on the underlying street network of each individual BA. Afterward, we set the relative gap between the incumbent solution and the best viable solution to 0% to obtain the optimal solution in each instance. We solved all instances in Julia 1.7 using JuMP 0.22 (Iain et al. 2017) and Gurobi 9.5 aborting the calculation of each instance after 21,600 s. Besides the variation of $|\mathcal{J}|$, we varied $|\mathcal{I}|$ and p. We did not restrict the maximum feasible driving time $\bar{d}$, thus assuming that each BA j can be assigned to each BA i. Furthermore, we solved each instance with our new constraints from the sets C1, C2 and C3 as well as without our new constraints (C0). We display the results in Table 1.

Summarized, $\mathcal{I}$ and $\mathcal{J}$ influence the calculation time in the same direction: larger sets lead to the longer computations. The change caused by a larger number of centers p is not as straightforward. Generally, it seems as if lower p values are more difficult to solve. This can be observed independent of the number of BAs or the different contiguity and compactness constraints. Our new contiguity and compactness constraints increase the calculation time except for one outlier: the stronger the compactness formulation, the higher the calculation time. The outlier occurred with 2040 BAs and potential locations, 3 centers and the constraint set C2. It is the only instance we couldn't solve within 6 h. If we ignore the outlier, our constraint sets C1, C2 and C3 increase the computation time by a factor of 3.08, 3.38 and 6.73 respectively when compared to the basic problem without contiguity and compactness constraints (C0). Thus, we conclude that our model can be applied to real problems without a heuristic despite the additional computation time caused by our new constraint sets, as long as the problem size is not too large.

4.5 Emergency service operation simulation

We cannot properly judge the quality of our solution based on our model alone, as it is an abstraction of the actual world. It incorporates neither queuing times, priorities and missing cars nor inter-district dispatches. One option would be the evaluation with real-world experiments, but this could lead to issues of safety, risks, and costs (Zhang and Brown 2013). Real scenarios are subject to uncertainties, dynamic

Table 1 Benchmark: different instances and their computation time (seconds)

$\lvert\mathcal{J}\rvert$	p	Ratio of BAs usable as department location ($\lvert\mathcal{I}\rvert/\lvert\mathcal{J}\rvert$)							
		50%				100%			
		Set of constraints				Set of constraints			
		C0[a]	C1[b]	C2[c]	C3[d]	C0[a]	C1[b]	C2[c]	C3[d]
510	3	4	9	20	20	12	37	53	71
	6	3	8	25	20	10	24	29	50
	9	6	11	23	25	8	16	19	46
	12	2	6	18	53	23	42	57	136
	15	4	7	18	28	7	16	25	58
	18	4	7	24	56	7	13	20	59
	21	3	5	23	51	15	31	45	75
	24	2	4	23	49	14	22	27	114
	27	3	6	26	44	9	18	21	148
	30	2	6	26	50	9	15	20	166
1008	3	26	100	145	170	59	499	299	413
	6	19	91	119	134	100	374	361	418
	9	19	74	107	125	97	209	225	681
	12	26	79	177	269	206	424	736	1,716
	15	12	44	57	209	75	155	193	234
	18	12	46	56	142	93	190	257	371
	21	18	63	74	343	72	141	181	311
	24	27	64	104	241	105	192	336	312
	27	35	126	123	240	81	146	156	522
	30	22	120	83	130	77	133	192	759
1508	3	57	264	449	539	247	1297	917	1,369
	6	86	281	417	362	143	822	721	1,011
	9	115	259	275	343	292	990	932	2,772
	12	98	219	255	832	479	1,494	1,766	4,988
	15	81	172	244	771	297	906	966	2,178
	18	77	154	227	751	335	690	794	917
	21	148	245	267	826	235	386	461	579
	24	87	169	259	1,466	280	490	557	1,337
	27	86	152	290	633	339	565	558	1,842
	30	62	121	227	435	305	439	547	901

Table 1 (continued)

$\|\mathcal{J}\|$	p	Ratio of BAs usable as department location ($\|\mathcal{I}\|/\|\mathcal{J}\|$)							
		50%				100%			
		Set of constraints				Set of constraints			
		C0[a]	C1[b]	C2[c]	C3[d]	C0[a]	C1[b]	C2[c]	C3[d]
2040	3	163	606	1,229	1,837	444	3,204	21,600[e]	9,821
	6	251	2,059	1,456	1,356	501	3,563	3,191	3,835
	9	818	1,215	1,896	1,299	869	4,837	3,649	8,113
	12	198	768	494	517	871	4,625	6,728	9,845
	15	266	665	938	1,395	564	2,337	2,508	5,989
	18	249	498	747	1,278	829	1,673	1,868	2,624
	21	186	806	465	1,650	686	1,686	2,504	3,444
	24	161	344	424	1,948	657	1,065	1,237	2,118
	27	363	1,018	801	2,525	723	1,234	1,339	2,961
	30	156	299	587	2,787	748	1,438	1,932	2,017

The computation time in seconds excludes the model-building time and displays the time Gurobi 9.5 needed to solve the instances with a relative gap of 0% between the incumbent solution and the best viable solution. All computations were done on an 8-Core AMD Ryzen 5700 G CPU with 32 GB RAM

[a]No contiguity and compactness constraints

[b]Contiguity constraint (9)

[c]Contiguity and compactness constraints (10) and (11)

[d]Contiguity and compactness constraints (10), (12) and (13). [e]No optimal solution could be found within 21,600 s

interactions and interdependencies between variables. It is difficult to create constant situations for experiments to derive meaningful conclusions (Miller and Knoppers 1972). To address these issues, we evaluated the resulting districts with an emergency service operation simulation. We wrote our simulation in Julia, a high-level programming language with an emphasis on speed (Bezanson et al. 2017). It allows for a fast simulation of real-world incidents, allowing us to simulate around 20,000 incidents per second on each CPU Core. Its structure is influenced by the emergency service simulations of Kolesar and Walker (1975), Kern (1989), Zhang and Brown (2013), Mayorga et al. (2013) and Dunnett et al. (2019) and augments these approaches. Our simulation is flexible and has various adjustable parameters to specify the dispatching rules. Among other things, it accounts for the variability of spatial and temporal incident patterns, multiple vehicle dispatch, shifts, priority handling, rush hours, exchanges between districts, variable driving times and follow-ups after incidents.

Algorithm 1 Pseudocode of the Simulation

```
for each minute in the observed timeframe do
    add all new CFS to a separate queue for each priority;
    for each priority do
        for each district do
            for each CFS in current district from queue of the current priority do
                if dispatching rules and available vehicles allow a dispatch then
                    dispatch min{available vehicles, requested vehicles};
                    subtract the dispatched vehicles from the requested vehicles of the CFS;
                    lower the available vehicles of the department until CFS is served;
                    if dispatched vehicle is the first vehicle assigned to CFS then
                        save the follow-up time caused by the CFS;
                    end
                    if CFS has no requested vehicles left then
                        remove CFS from the queue;
                    end
                else
                    continue with the next incident in queue;
                end
            end
        end
        for each CFS in the queue of the current priority do
            if CFS has not been assigned at least one vehicle then
                if dispatching rules and available vehicles allow a dispatch from a different
                  department then
                    determine the best department according to the dispatching rules;
                    dispatch one vehicle from the selected department;
                    subtract the dispatched vehicle from the requested vehicles of the CFS;
                    lower the available vehicles of the department until CFS is served;
                    if dispatched vehicle is the first vehicle assigned to CFS then
                        save the follow-up time caused by the CFS;
                    end
                    if CFS has no requested vehicles left then
                        remove CFS from the queue;
                    end
                else
                    continue with the next incident in queue;
                end
            end
        end
    end
    for each district do
        if district has available vehicles then
            if the follow-up time after CFSs has not been completed yet then
                assign available vehicles to work on the follow-up of their incidents;
            end
        end
    end
end
```

The basic idea is described in the pseudocode of Algorithm 1. Our simulation runs over a specified time horizon and serves the historical incidents during that time. We start with a given number of vehicles per district, which changes over the course of a weekly shift pattern. Each minute, we allocate all new incidents to queues. Afterward, we allocate the number of requested cars to each incident based on the predefined dispatching rules. After vehicles have served an incident, the bound resources enter the pool of available vehicles again, either to be dispatched anew or to work on the deployment follow-up. Besides the simulation of district layouts, our simulation allows for experiments with dispatching rules, shift patterns and personnel resource allocations. Several important factors to our case study, such as the consideration of variable driving times and the computation of the number of vehicles per district, are described in the next section. But to keep our work within a reasonable length, we refrain from describing other factors, such as the dispatching rules, in more detail. For interested readers who would like to know more, we published our framework together with exemplary data and further explanations about the simulation and the dispatching rules on GitHub.[1]

5 Case studies

5.1 Data and parameters

In this section, we apply our framework to improve the districting of two confidential regions in Germany and Belgium. To solve our research questions, the authorities granted us access to two classified incident data sets. The first data set consisted of all CFS in a confidential region in Germany from 2015 to 2019. The second data set included all CFS in Belgium from 2019 to 2020. Because of the confidentiality of the data, incident locations were partly aggregated at street level. All incidents encountered by patrol cars and incidents reported in retrospect were not included in the data sets.

With this foundation, we set up a hexagonal representation of both areas. Except for the driving time d_{ij} and the Euclidean distances a_{ij}, we could derive all necessary parameters for each BA j from the police data sets or use approximations in consultation with the police authorities. For the computation of the Euclidean distances a_{ij} necessary for our contiguity and compactness constrains, we use the original centroids of the BAs. The driving time d_{ij} represents the time to get from the centroid of i to the centroid of j on the underlying street network. We thus moved the centroids to the nearest major street intersection to increase the accuracy of the driving times, as proposed by Shen et al. (1995). In addition, this ensures that remote areas with few streets are still reachable if we impose the maximal allowed driving time $\bar{d}$. The street network in Germany was based on OpenStreetMap Data (Geofabrik GmbH and OpenStreetMap Contributors 2020). TomTom[2] supplied the street network in Belgium. For the computations, we used Dijkstra's algorithm and took the speed limits of all roads into account. Although the responders are not legally bound by speed limits, they serve as an indicator of the road conditions. For example, police cars within cities usually cannot drive at 90 or 100 km/h, as they must drive carefully. On country roads they can drive much faster, e.g., at 120 or 130 km/h - but even here they must look after animals crossing the street. Thus, they can never really drive unconstrained (except on highways). Now, d_{ij} represents the driving time without congestion, traffic lights, construction sites or other obstacles. In the real world, these conditions have a substantial influence on the driving time, especially in urban areas (Zhang and Brown 2014). With this in mind, we introduce the traffic

1 https://github.com/beyondsimulations/emergencyservices.

2 ©TomTom, https://www.tomtom.com/.

flow parameter r_t to enhance and legitimize our driving time. It is a multiplier for each week-hour $t \in \mathcal{T}$ and describes the additional time required to move between two locations during a "typical week". In accordance with the available data, we selected a week during springtime that was free of public holidays. To determine r_t, we utilized the Google Maps Distance Matrix API[3] to anticipate two driving times for eight different routes between the BAs l and k with $l \in \mathcal{I}$ and $k \in \mathcal{J}$ for each week-hour t during the selected week. The first was an optimistic driving time d^o_{klt} and the second was the best guess of the driving time d^b_{klt} generated by the Google Maps forecasting model. We didn't include the pessimistic forecast, as police vehicles under time constraints move faster than regular vehicles. Furthermore, we estimated that police vehicles arrive on average between d^o_{klt} and d^b_{klt}.

To determine r_t, we first divided the forecasted driving times of the API through our calculated driving time d_{kl} from the corresponding route for each week-hour, and averaged the quotients over all routes as expressed in (14).

$$r_t = \frac{\sum_k \sum_l (d^o_{klt} + d^b_{klt})}{\sum_k \sum_l 2d_{kl}} \quad \forall t \in \mathcal{T} \tag{14}$$

Next, we used r_t to derive the improved mean driving time d_{ijt} for all routes with (15).

$$d_{ijt} = r_t \times d_{ij} \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}, \forall t \in \mathcal{T} \tag{15}$$

Ideally, we would have forecasted the driving times for all routes between our BAs for each week-hour with the API, but this was economically unfeasible due to the cost of each API request.

5.2 Optimization design

Our optimization has an identical structure in both case studies. Because of the partly aggregated incident location data and a strong deviation of requested vehicles between the BAs, we settled for hexagons with a diameter of 750 ms in both case studies, which resulted in 1596 BAs in Germany and 1233 BAs in Belgium. We mapped all natural barriers in the adjacency matrix. In addition, we had to account for the fact that police departments have to fulfill several elements of community demand rather than just CFS (McCabe 2017). We thus discarded all rural BAs without streets or with zero incidents and all BAs where it would not be suitable to locate a police department because of their geography from the set of the potential locations $\mathcal{I}$. Thereafter, we were left with 1166 potential department locations $\mathcal{I}$ in Germany and 914 potential locations $\mathcal{I}$ in Belgium. A summary is displayed in Table 2.

In Germany, we solved three distinct problems: G1, G2, G3. First, we applied our framework to the optimization of the district borders with the current department

[3] ©Google, https://cloud.google.com/maps-platform/.

Table 2 Optimization summary

	Germany	Belgium
Current departments[a]	20	2
Data set horizon	5 years	2 years
Potential locations $\lvert\mathcal{I}\rvert$	1,166	914
Basic areas $\lvert\mathcal{J}\rvert$	1,596	1,233
Hexagonal diameter	750 ms	750 ms
Incidents[a]	1,800,000	50,000
Max. driving time $\bar{d}$	18 min	26 min
Max. supporting drive $\bar{d}^q$	18 min	32 min
Min. supporting centers	2 centers	1 center
Priority weights w_p	$\left(4\,3\,2\,1\right)^\intercal$	$\left(5\,4\,3\,2\,1\right)^\intercal$

[a]Approximate value due to confidentiality

locations (G1). Next, we allowed the exchange of the location of one police department (G2). Then, we allowed the reallocation of all police departments (G3). We refer to the current layout in Germany as G0. In Belgium, we solved the three separate problems B1, B2 and B3, and focused on different objectives, as the authorities were interested in the optimal location of a new police department. First, we optimized the district borders while fixing the present two department locations (B1). Thereafter, we kept the two present department locations and allowed the model to allocate a third police department freely (B2). Last, we granted the free allocation of three police departments (B3). To refer to the present layout in Belgium, we use the abbreviation B0. We examined all six problems G1, G2, G3, B1, B2 and B3 for each contiguity variation C0, C1, C2 and C3 from Sect. 4.3.

Please note that the datasets analyzed during the case study are not publicly available due to the strict confidentiality of the information. However, we provide sample data sets on GitHub that are close to the structure of the real data.

5.3 Simulation design

We simulated the police service over the entire time horizon of each historic data set to assess the new district layouts. Naturally, the driving time of each incident response exhibits variability in the actual world. Moreover, the variance of the driving time changes over the course of each week, as a lower traffic volume, especially at night, contributes to better predictability. To account for this, we replaced r_t in our simulation with a normally distributed variable $q_t \sim N(r_t, \sigma_t)$. To determine σ_t, we used the predictions from the Google Maps Distance Matrix API. Furthermore, we estimated that the traffic flow parameter q_t lies between

$$r_t^b = \frac{\sum_k \sum_l d_{klt}^b}{\sum_k \sum_l d_{kl}} \quad \forall t \in \mathcal{T} \tag{16}$$

and

$$r_t^o = \frac{\sum_k \sum_l d_{klt}^o}{\sum_k \sum_l d_{kl}} \quad \forall t \in \mathcal{T} \tag{17}$$

in 90% of all incidents. Thus, the stochasticity of the traffic and other unforeseen circumstances in the corresponding week-hours were incorporated into the simulation.

In addition, we had to define new plans for the shift pattern in each police department, as our new layouts influenced the required personnel resources. We used a heuristic based on safety stocks to plan the resources based on the accrued workload over the time horizon of each data set. $\mathcal{H}$ was the set of all departments and $\mathcal{K}$ the set of all week-hours and $\mathcal{K}^S$ was a partition of the week-hours $\mathcal{K}$ into shifts. First, we calculated the average workload μ_{hk} and standard deviation σ_{hk} of each department $h \in \mathcal{H}$ in each week-hour $k \in \mathcal{K}$ based on the district layout. Next, we used equation (18) with the inverse normal distribution to determine the adjusted weekly time demand b_{hk}.

$$b_{hk} = F_{Z_{hk}}^{-1}(\alpha_s) \text{ with } Z_{hk} \sim N(\mu_{hk}, \sigma_{hk}^2) \tag{18}$$

In both case studies, we applied the service factor $\alpha_s = 0.90$, as departments under pressure could request vehicles from surrounding districts. We then used b_{hk} to determine the highest demand $\tau_{hs} = \max_{k \in s} b_{hk} \forall s \in \mathcal{K}^S$ of each location h and shift $s \in \mathcal{K}^S$. Afterward, we assigned the total weekly available workforce of the police jurisdiction to each shift, depending on the highest demand τ_{hs} of each shift. In addition, we accounted for the minimum workforce that had to be assigned to each district in each shift. Thus, we determined the shift plan in each department a posteriori based on the incidents we later used in our simulation. We fixed the resulting weekly shift plans over the whole time horizon of the simulation. Our focus was neither shift planning nor forecasting; thus we deemed these assumptions acceptable as they allowed us to test our district layouts with good shift plans. The dispatching rules were constant over the entire simulation and were defined in accordance with the police administrators. To ensure stable results, we simulated each layout 100 times and reported the average of each criterion. Both optimization model and simulation ran on the same workstation as before.

5.4 Results in Germany

Table 3 displays the computation time for all six problems and their variations from our two case studies. Except for the outlier problem G3 with contiguity variation C3, we could solve all problems in Germany within 6 min.

To evaluate the improvements of our new layouts in Germany, we compared selected criteria from our simulation with the present layout G0. In all simulations, we employed the same dispatching rules and resource parameters. We display the results for Germany in Table 4. We focus on the districts from the constraint sets C1, C2 and C3, as the districts without these constraints were non-contiguous. Because of the confidentiality of the data, we can only show the relative change to the present layout G0.

Table 3 Computation time for all problem variations in Germany and Belgium (seconds)

	Constraint set			
	C0	C1	C2	C3
G1	0.20	0.68	0.77	5.14
G2	35.81	135.86	121.25	197.27
G3	70.12	224.25	347.02	3167.83
B1	0.03	0.05	0.06	0.06
B2	34.89	79.35	97.59	119.45
B3	57.05	124.01	142.41	193.62

The computation time in seconds excludes the model-building time and displays the time Gurobi 9.5 needed to solve the instances with a relative gap of 0% between the incumbent solution and the best viable solution. All computations were done on an 8-core AMD Ryzen 5700 G CPU with 32 GB RAM

Table 4 Germany: results from the simulation (%)

		DP	RS	OD	TD	EX
G1[a]	C0[d]	−10.59	−6.41	−4.57	3.28	−3.02
	C1[e]	−8.82	−5.58	−4.15	2.92	−2.47
	C2[f]	−9.27	−5.39	−3.67	2.54	−1.75
	C3[g]	−10.53	−6.00	−4.00	2.75	−2.95
G2[b]	C0[d]	−12.67	−7.50	−5.22	3.62	−3.59
	C1[e]	−10.78	−6.47	−4.56	3.05	−2.32
	C2[f]	−9.22	−5.74	−4.20	2.86	−1.69
	C3[g]	−10.56	−6.11	−4.15	2.76	−2.07
G3[c]	C0[d]	−20.24	−13.29	−10.22	6.76	0.21
	C1[e]	−21.16	−13.60	−10.27	6.72	0.39
	C2[f]	−19.98	−12.98	−9.89	6.46	0.93
	C3[g]	−22.56	−14.52	−10.98	7.37	−4.22

The table displays the relative change of each criterion compared to the current layout G0. The compared criteria are the dispatch time (DP), the response time (RS), the overall driving time (OD), the time at the department (TD) and the exchange ratio (EX)

[a]Current department locations with optimal district borders

[b]Optimal solution with one relocatable department

[c]Optimal solution in case all departments are relocatable

[d]No contiguity and compactness constraints

[e]Contiguity constraint (9)

[f]Contiguity and compactness constraints (10) and (11)

[g]Contiguity and compactness constraints (10), (12) and (13)

The first criterion represents the improvement of the dispatch time. It shows the average time until the first responding vehicle is assigned to an incident. In our simulation, this time could be decreased by up to 10.53% while using the present department locations in problem G1 with the constraints from C3, and up to 22.56% if we allowed the reallocation of all police departments in G3 while using C3. The second criterion expresses the difference of the average response time of the first responder. Given the established department locations in problem G1, we could shorten it by up to 6.00% with constraint set C3. If we allowed the reallocation of one department in problem G2, the best reduction amounted to 6.47% with C1. The best improvement by 14.52% was achievable in problem G3 with C3. The third criterion shows the difference of the overall driving time in the simulation. The optimization lowered the driving time on average by 4.10% in G1, 4.53% in G2, and 10.34% in G3. While this might seem small at first, an improvement of 1% is roughly equivalent to a weekly reduction of the driving time by 17 h over all departments. The fourth criterion displays the relative change in the time vehicles spent at the department. At the department, vehicles—and thereby the officers—are not assigned to specific tasks. Thus, they can react immediately in case of a call for service. Time spent at the department increased in G1, but the improvement was even greater in G2 and G3. The last criterion covers the exchange rate. During our simulation, an exchange took place, if a district did not have enough resources to dispatch a vehicle. In this case, other districts had to dispatch their cars to help. Our new districts mostly lowered the exchange rate, the best improvement by 4.22% was achieved in problem G3 with C3.

A closer analysis of the contiguity variants shows we could achieve the strongest improvements in G1 with C0, although the gains of C3 were close behind. Furthermore, we could produce the best improvements in G2 with C0 and in G3 with C3. The discrepancy of all criteria between the variants C0, C1, C2 and C3 of each problem remained within 3.5%. The driving time of each simulated problem instance with C0 was the smallest for G1 and G2, although the resulting districts were neither contiguous nor visually compact. In G3, the driving time was the lowest with the constraints from C3. A further analysis showed that the main reason for the reduction was the greater compactness of the districts that allowed for better inter-district exchanges under heavy load. But to benefit from the stricter contiguity constraints, the department locations had to be unconstrained to expand the permissible solution range.

A look at the district visualization in a geographic information system reveals the increased compactness. Note, that we only present the layouts where all departments were relocatable and that we have altered the covered area to protect the confidentiality of the region. This results in some districts, where the district center is outside the displayed area. The district from problem G3 with C0 is displayed in Fig. 2 and G3 with C1 is shown in Fig. 3. Figure 4 displays the best layout G3 with C2, and Fig. 5 illustrates the best layout of problem G3 with C3. It is apparent that the districts in Figs. 4 and 5 exhibit a visibly higher compactness than the districts in Figs. 3 and in 2. Equally important, Fig. 5 shows that the departments are distributed more evenly across the map when compared to the other district layouts.

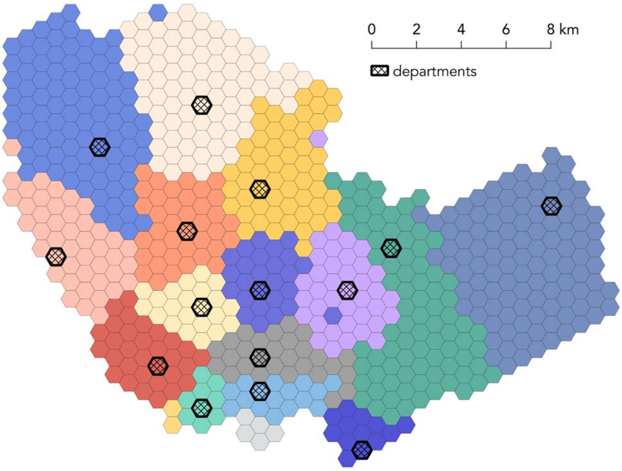


Fig. 2 Optimal solution if all departments are relocatable (G3) without contiguity constraints (C0). Response time reduced by 13.29%

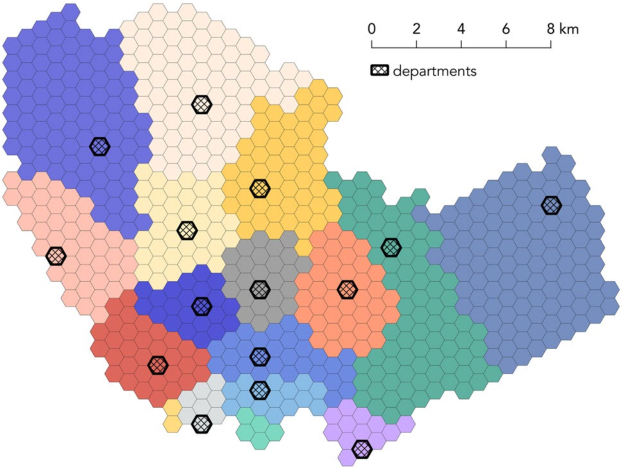


Fig. 3 Optimal solution if all departments are relocatable (G3) under contiguity constraint (9) (C1). Response time reduced by 13.60%

In addition, the new districts improved the number of unanswered CFS. We classified incidents as unanswered, if they couldn't be served within 6 h. Except for a few outliers, they belonged to the lowest priority. They were queued and skipped because the vehicles had to attend incidents with a higher priority. After the optimization, the number of missed incidents compared to the present layout was lower in seven of nine instances.

During our work, we also examined the influence of balancing the workloads per department enforced by additional constraints, as the minimum workforce that

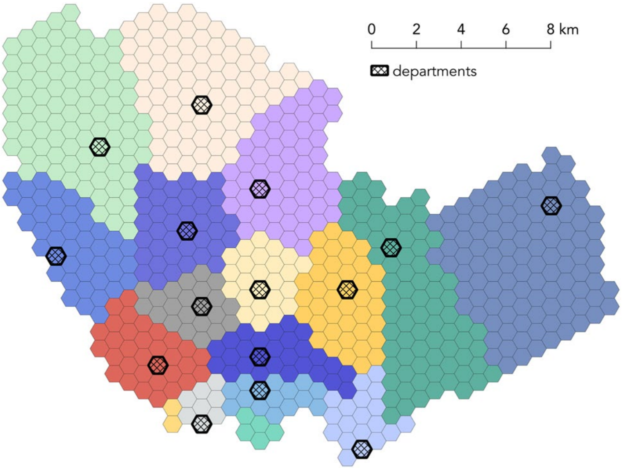


Fig. 4 Optimal solution if all departments are relocatable (G3) under contiguity and compactness constraints (10) and (11) (C2). Response time reduced by 12.98%

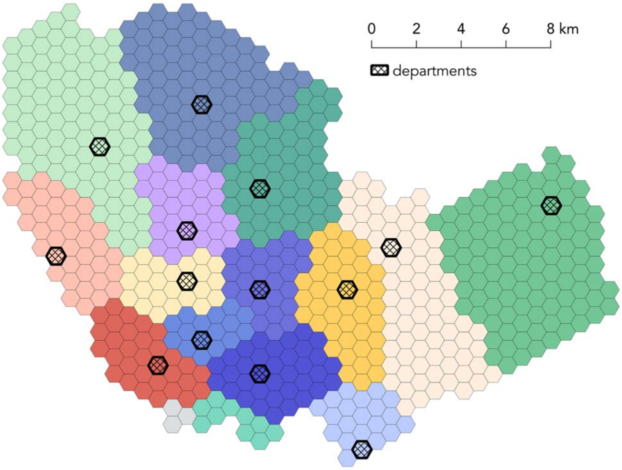


Fig. 5 Optimal solution if all departments are relocatable (G3) under contiguity and compactness constraints (10), (12) and (13) (C3). Response time reduced by 14.52%

has to be ensured in each department might lead to unbalanced resource demands during low-workload periods. Our idea was that balanced workloads might contribute to lower dispatch times and therefore to lower response times because of better resource distribution. But there is a trade-off between a restriction to the workload and a reduction of the driving time because further restrictions lower the permissible solution range. Unfortunately, our simulation demonstrated that the longer driving time negated any improvements arising from a better workload distribution.

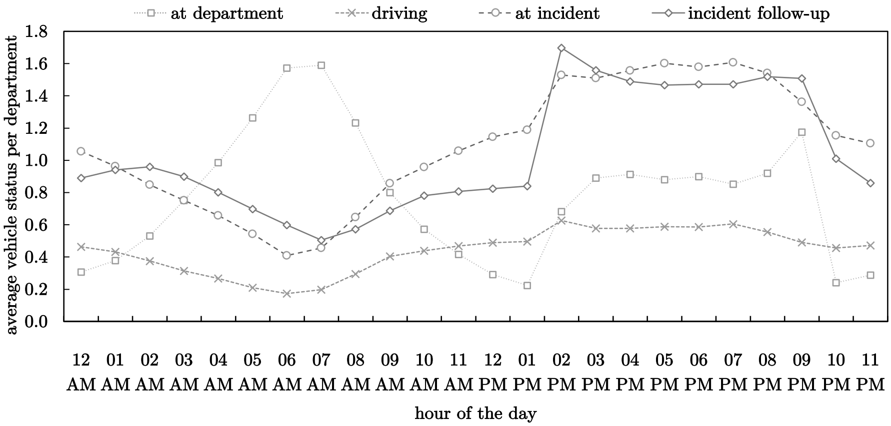


Fig. 6 Current layout in Germany (G0): average number of vehicles per status over the course of 24 h

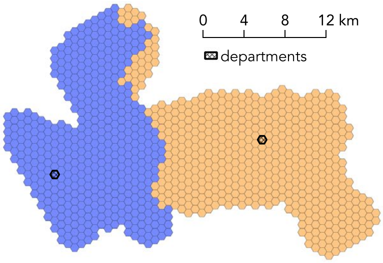
In addition, our simulation showed that the shift pattern, the resource distribution and the dispatching rules had a strong impact on the criteria. Figure 6 illustrates the average vehicle status per department for each hour of the day over the whole timeframe of the simulation in G0 under an artificial shift pattern. In this example, each shift was 8 h long and new shifts began at 6 a.m., 2 p.m. and 10 p.m. every day. Among other things, the figure shows a spike in vehicles 'at department' between 4 a.m. and 9 a.m. This spike arose because of suboptimal shift changes, which contributed to wasted personnel resources: the assigned vehicles of the shifts between 6 a.m. and 2 p.m. had to respond to a rising incident volume each hour. At 6 a.m. the departments had around 1.6 vehicles without duty, while the system was under heavy load with only 0.2–0.4 free vehicles between 11 a.m. and 2 p.m. A similar reversed pattern developed in the shifts from 10 p.m. to 6 a.m. Intensified through the spatio-temporal variability of the incidents and the variability of the driving time, this led to many inadequate responses in our simulation. Overall, the officers spent only 23.27% of the time at the department without a task. This is critical because a part of this time did not serve as a safety buffer but emerged from suboptimal shift plans.

In conclusion, we could improve all five criteria in Germany through our new district layouts. Nevertheless, we have to keep in mind that other factors influence the efficiency and effectiveness of service operations as well.

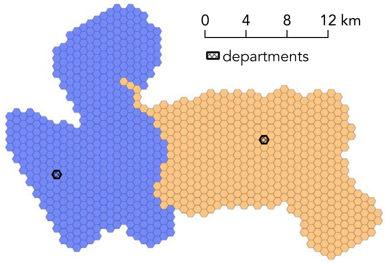
5.5 Results in Belgium

In our case study in Belgium, we applied the same framework as in Germany, but we adjusted the dispatching rules to account for the structure of the spatial incident distribution. We could solve all optimization instances in Belgium within 4 min, as disclosed in Table 3.

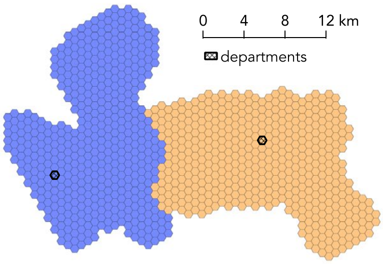
The district layouts of the contiguity variations C1, C2 and C3 in Belgium were nearly identical in each problem. Hence, the differences between the



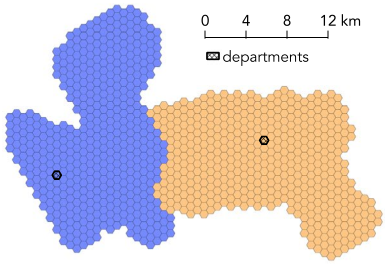
(a) No constraints (C0). Response time reduced by 6.08 %.



(b) Contiguity constraint (9) (C1). Response time reduced by 3.28 %.



(c) Contiguity and compactness constraints (10) and (11) (C2). Response time reduced by 3.34 %.



(d) Contiguity and compactness constraints (10), (12) and (13) (C3). Response time reduced by 3.30 %.

Fig. 7 Current department locations with optimal district borders (B1) in Belgium

contiguity variations of each problem were much smaller than in Germany. We illustrate this in Fig. 7 for B1. Note that we have again altered the covered area to protect the confidentiality of the region. These lower deviations can be explained through the lower number of police departments and the geography of the region. In contrast to Germany, the region in Belgium is not a metropolitan area but a cluster of towns with rural areas in between. The district borders of the layouts run through rural regions without sizeable incident volume variations. Nevertheless, our contiguity variations helped to ensure contiguity and to enhance the compactness. Figure 7a displays the optimal solution without our new constraints for the current department locations. Clearly, the solution is neither contiguous nor compact. Figure 7b illustrates that the contiguous solution from C1 results in a non-compact distortion that the identical solution from the more restricted contiguity variations C2 in (c) and C3 in (d) prevented.

The simulated improvements of our layouts compared with the present layout B0 are displayed in Table 5. The dispatch time in the three instances in B1 was greater than in the present district layout B0. In particular, assigning three departments instead of two, as in B2 and B3, contributed to an increase in the dispatch time and a sharp rise in the exchange ratio. We expected this, as equal personnel

Table 5 Belgium: results from the simulation (%)

		DP	RS	OD	TD	EX
B1[a]	C0[d]	−6.08	−4.64	−3.31	1.17	−14.34
	C1[e]	−3.28	−3.29	−3.30	1.23	−14.86
	C2[f]	−3.34	−3.32	−3.31	1.24	−14.70
	C3[g]	−3.30	−3.30	−3.30	1.24	−14.71
B2[b]	C0[d]	0.95	−4.41	−9.36	3.46	145.05
	C1[e]	0.97	−4.41	−9.38	3.47	144.76
	C2[f]	0.93	−4.43	−9.38	3.47	144.85
	C3[g]	1.18	−4.30	−9.36	3.45	145.44
B3[c]	C0[d]	−5.81	−8.47	−10.93	3.94	145.05
	C1[e]	−5.86	−8.50	−10.94	3.93	145.12
	C2[f]	−5.19	−8.11	−10.81	3.94	147.75
	C3[g]	−6.18	−8.66	−10.95	3.94	145.06

The table displays the relative change of each criterion compared to the current layout B0. The compared criteria are the dispatch time (DP), the response time (RS), the overall driving time (OD), the time at the department (TD) and the exchange ratio (EX)

[a]Current department locations with optimal district borders

[b]One additional department location with optimal districts

[c]Free allocation of three departments with optimal districts

[d]No contiguity and compactness constraints

[e]Contiguity constraint (9)

[f]Contiguity and compactness constraints (10) and (11)

[g]Contiguity and compactness constraints (10), (12) and (13)

resources had to be shared among more departments, which led to a longer dispatch time under heavy workloads for incidents with a low priority and more exchanges. Nevertheless, the response time was shorter in B2 and B3 because of the strong reduction in driving time. An exception was the non-contiguous district layout from Fig. 7a, which response time was only beat by the layouts in B3. A second benefit of the additional location was more time spent at the department.

In contrast to Germany, our new layouts in Belgium did not improve each criterion. However, our framework showed that a new department can reduce the response time and the driving time significantly. Whether this justifies the drawbacks in other criteria has to be decided by the authorities.

6 Discussion

The results of our case studies are promising, as the simulation demonstrated that our new framework can improve several criteria. A modification of our optimization problem could be applied to the key problem of the PDP: the planning of patrol beats within districts. Besides police services, we could employ an altered version of our framework for fire and ambulance services, as the structure is closely related. Moreover, if requested, we could implement extensions with restrictions on either the workload deviation or the department capacities. These measures would benefit from our newly proposed contiguity constraints, as the likelihood of distortions occurring in the district shapes increases with new constraints. In addition, our newly proposed constraints could be applied to other TDPs to achieve contiguity and compactness within reasonable computation times. Furthermore, additional research can address the current limitation of these constraints to a hexagonal BA structure to support BAs of all shapes.

But our framework has some limitations. First, we assume that the future incident pattern reflects the historic pattern. Moreover, in some instances the district layouts without contiguity and compactness constraints (C0) featured the lowest response time. One reason is that our simulation doesn't include patrol cars as backup because this wasn't a research goal in either case study. Nevertheless, these backups might improve the simulated criteria of the new districts resulting from C2 and C3, as these districts exhibit an improved compactness. Police administrators must decide whether giving up these conditions might be worthwhile considering the potentially lower response time, or whether we should expand our simulation in a future study to include these backups. In addition, the resulting districts don't follow municipal boundaries, as we used hexagons as BAs. Last, spatio-temporal incident patterns change and implementing new district layouts and department locations might influence and advance this process.

Apart from that, our work revealed several other influences on the response time. Smaller shift intervals, an adjustment of shift changes, a better capacity distribution and more personnel could lower the response time and improve other criteria. These influences could be investigated further in future research. To achieve this, we could replace the capacity heuristic with an optimization model, aiming to distribute the overall capacities, shift lengths and shift changes depending on the workload, the legal rules and the officers' preferences. In addition, we could improve the dispatching rules, as in our case studies dispatchers decide which vehicle to dispatch based on their best knowledge. Our simulation would support experiments with variations to the present rules, and additional research could be conducted to complement the dispatchers with a recommendation system, as proposed by Shen et al. (1995).

Above all, our work and our district layouts have to be questioned thoroughly and with care. We only used CFS for our optimization, and we didn't model other elements of community demand (McCabe 2017). Furthermore, certain crimes are more likely to be reported, and only recorded crimes become data. In addition, citizens in certain areas can be uncomfortable with the police and thus avoid reporting incidents. Hence, modeling districts and especially patrols beats after historic incidents without further inquiries can lead to blind spots (Vestby and Vestby 2021). Likewise, we cannot ignore that crime expresses discontent in the social system. Our framework does not address the roots of such discontent and solely deals with the surface of the problem. Community work and political approaches have to be addressed simultaneously for sustainable improvements (Peet 1975).

7 Conclusions

In this paper, we introduced a districting framework that helped to provide valuable information to police administrators in Germany and Belgium. The first part of the framework is based on a linear mixed-integer optimization that reduces the driving time of police vehicles using a novel combined approach to compactness and contiguity through a hexagonal grid. The problem with non-fixed centers is valuable in the design phase of a police system. Once the district centers have been built, a reoptimization of districts is an interesting managerial problem in its own right that should regularly be carried out, for instance to react to a change in the frequency of CFS. The second part is a simulation to evaluate the quality of the resulting layouts according to several criteria. The potential improvements are promising and followed from our newly introduced constraints. In Germany, our simulation showed a potential reduction of the response time by 6.00% through a redesign of the present district borders. A substantial response time reduction by 14.52% could be achieved through the reallocation of several departments. In Belgium, we could reduce the response time by up to 8.66% through the introduction of a third police department.

Different other criteria in both case studies could be improved as well. In conclusion, our framework is suited not only for redistricting, but it can also help with strategic planning decisions regarding new department locations. Moreover, our results show that police service administrators could benefit from the use of OR techniques in the planning of department locations and district layouts. The sooner they incorporate such techniques into the planning process, the more room for improvements.

Appendix: proof of *NP*-completeness

In the following, we establish *NP*-Completeness of the decision version of CPMFC. *NP*-hardness of the optimization version then follows trivially.

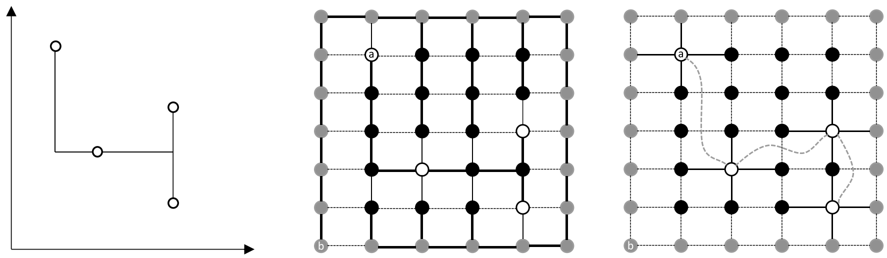
p-Median Contiguous Districting - decision version (CPMFC-d): Let $G = (\mathcal{J}, \mathcal{E})$ be an undirected graph and $\mathcal{I}^s \subseteq \mathcal{J}$ be a set of given district centers with cardinality P. Further, let $t : \mathcal{I}^s \times \mathcal{J} \rightarrow \mathbb{N}$ denote the total driving times between district centers and BAs. Find a partition of the nodes of G into P connected subgraphs U_i, with $\left|U_i \cap \{i\}\right| = 1 \ \forall i \in \mathcal{I}^s, U_i \cap U_{i'} = \emptyset \ \forall i, i' \in \mathcal{I}^s, i \neq i'$, such that $\sum_{i \in \mathcal{I}^s} \sum_{j \in U_i} t_{ij} \leq k$.

We will give a pseudo-polynomial reduction from the Rectilinear Steiner Tree Problem which is well known to be *NP*-complete in the strong sense (Garey and Johnson 1977) and is restated in the following:

Rectilinear Steiner Tree Problem (RSTP) Given a set A of points in the plane. Does there exist a tree structure which connects all points in A with vertical and horizontal unit line segments such that their total length is no larger than some constant k?

We are now ready to prove Theorem 1.

Proof of Theorem 1 It can be easily seen that CPMFC-d is in *NP* since the connectedness of the subgraphs and the value of the objective function itself can be



(a) Instance of RSTP (b) Instance of CPMFC (c) Road Network of CPMFC Instance

Fig. 8 Reduction to p-median districting from rectilinear steiner trees

determined in polynomial time in the number of vertices and edges of the graph. Further, Let *maxX*, *maxY*, *minX* and *minY* denote the maximum/ minimum *x*- or *y*-coordinate of all points in *A*, then consider a grid graph *G* with width of $maxX - minX + 2$ and height of $maxY - minY + 2$. For convenience, let the labels of the graph correspond to the coordinates in the plane such that the top left node in the grid graph corresponds to point $(minX - 1, maxY + 1)$ in the plane, the top right to $(maxX + 1, maxY + 1)$, bottom left to $(minX - 1, minY - 1)$ and bottom right to $(maxX + 1, minY - 1)$ and all other nodes correspondingly. Let the set *A* of points from the RSTP correspond to a set $A \subseteq G$ of nodes in the graph, and let the number of district centers be $P = 2$. Without loss of generality, we assume that no node in *A* is adjacent in the graph (should two nodes be adjacent we can just increase the precision in the translation between the points in the plane and the grid graph, e.g. by additionally considering the half-points in between any two points and relabel the nodes accordingly). Now, select an arbitrary node *a* in *A* as district center 1 and set $t_{aj} = 1 \; \forall j \in \mathcal{J} \setminus \{a\}$ and $t_{aa} = 0$. Let the second district center be located at node $b = (minX - 1, minY - 1)$ and set $t_{bj} = 0 \; \forall j \in \mathcal{J} \setminus A$ and $t_{bj} = k + 1 \; \forall j \in A$. Notice that any such instance of CPMFC-d can be generated from an instance of the RSTP in pseudo-polynomial time in the size of the input of RSTP.

Now, consider a solution to a YES-instance of RSTP and let *T* be the set of points that are connected by unit line segments in the tree structure. Notice that by necessity $A \subseteq T$ and $k = |T| - 1$. Consider the corresponding nodes in the grid graph and assign all nodes in *T* to center *a* and the remaining nodes to *b*, so that $U_a = T$ and $U_b = \mathcal{J} \setminus T$. Since *T* is a tree structure, the district assignment U_a is contiguous. Further, since it will hold for all points $(xy) \in T$ that $minX \leq x \leq maxX$ and $minY \leq y \leq maxY$, *T* has no loops and no two nodes in *A* are adjacent, also the set U_b is contiguous. As $t_{aj} = 1 \; \forall j \in \mathcal{J} \setminus \{a\}$ and $t_{aa} = 0$ and further *T* contains $k + 1$ nodes including *a* it follows that $\sum_{j \in U_a} t_{aj} = k$. Finally, since $t_{bj} = 0 \; \forall j \in \mathcal{J} \setminus A$ it follows that $\sum_{j \in U_b} t_{bj} = 0$ and thus the answer to the corresponding instance of the CPMFC-d will be YES.

In turn, consider a solution of a YES-instance of CPMFC-d and let *T* be the set of districts assigned to the first district center $T = U_a$. Notice that since $t_{aj} = 1 \; \forall j \in \mathcal{J} \setminus \{a\}$ it holds that $|T| \leq k + 1$ and that $A \subset T$. As *T* is contiguous, we can now translate the nodes in T to a set of adjacent points in the plane and since $|T|$ adjacent integer points can be connected by $|T| - 1$ unit line segments, the answer to the corresponding instance of the RSTP will be YES, which completes the proof. □

Figure 8 shows a YES-instance of the RLSP and the corresponding CPMFC-d instance. Furthermore, it contains an example of a road network that would give rise to the total driving times employed in the proof as long as driving times can depend on traffic intensity and/or ease of connection and demand intensities are considered. To see that, assume that the dashed rectilinear connections between nodes of the grid graph constitute very fast lanes without congestion, such that driving times are negligible. Furthermore, driving times for each node to itself are also negligible. The dashed curved connections are fast lanes with a driving time of $\frac{1}{K}$ that exclusively connect points in A (say by a set of bridges or tunnels). Finally, the bold connections to and from the points in A are congested roads with a driving time of 1. It thus follows that $d_{aj} = \frac{1}{K}\ \forall j \in A \backslash a$, $d_{aa} = 0$ and $d_{aj} = 1\ \forall j \in \mathcal{J} \backslash A$. For center b driving times amount to $d_{bj} = 1 + \frac{1}{K}\ \forall j \in A$ and $d_{bj} = 0\ \forall j \in \mathcal{J} \backslash A$. Let e_j denote the demand intensities and let the nodes in A be of high demand $e_j = k\ \forall j \in A$ and the rest of low demand intensity $a_j = 1\ \forall j \in \mathcal{J} \backslash A$. We can then easily check that this leads to the total driving times employed in the proof.

This demonstrates that unless sufficiently different velocities and/or demand intensities can be ruled out, exact algorithms are unlikely to guarantee even (pseudo-) polynomial worst-case performance for contiguous districting problems under the p-median objective.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Altman M (1997) Is automation the answer: the computational complexity of automated redistricting. Rutgers Comput Law Technol J 23:81–142

Badri MA, Mortagy AK, Alsayed CA (1998) A multi-objective model for locating fire stations. Eur J Oper Res 110(2):243–260. https://doi.org/10.1016/s0377-2217(97)00247-6

Bezanson J, Edelman A, Karpinski S et al (2017) Julia: a fresh approach to numerical computing. SIAM Rev 59(1):65–98. https://doi.org/10.1137/141000671

Bodily SE (1978) Police sector design incorporating preferences of interest groups for equality and efficiency. Manage Sci 24(12):1301–1313. https://doi.org/10.1287/mnsc.24.12.1301

Bruce C (2009) Districting and resource allocation: a question of balance. Q Bull Appl Geogr Study Crime Public Saf 1(4):1–3

Camacho-Collados M, Liberatore F (2015) A decision support system for predictive police patrolling. Decis Support Syst 75:25–37. https://doi.org/10.1016/j.dss.2015.04.012

Chaiken JM, Dormont P (1978) A patrol car allocation model: capabilities and algorithms. Manage Sci 24(12):1291–1300. https://doi.org/10.2307/2630694
Cheung CY, Yoon H, Chow AH (2015) Optimization of police facility deployment with a case study in greater London area. J Facil Manag 13(3):229–243. https://doi.org/10.1108/JFM-05-2014-0014
Cordone R, Maffioli F (2004) On the complexity of graph tree partition problems. Discret Appl Math 134(1–3):51–65. https://doi.org/10.1016/S0166-218X(03)00340-8
Curtin KM, Qui F, Hayslett-McCall K, et al (2005) Integrating GIS and maximal covering models to determine optimal police patrol areas. In: Wang F (ed) Geographic information systems and crime analysis. IGI Global, pp. 214–235, https://doi.org/10.4018/978-1-59140-453-8.ch013
Curtin KM, Hayslett-McCall K, Qiu F (2010) Determining optimal police patrol areas with maximal covering and backup covering location models. Netw Spat Econ 10(1):125–145. https://doi.org/10.1007/s11067-007-9035-6
Dunnett S, Leigh J, Jackson L (2019) Optimising police dispatch for incident response in real time. J Oper Res Soc 70(2):269–279. https://doi.org/10.1080/01605682.2018.1434401
D'Amico SJ, Wang SJ, Batta R et al (2002) A simulated annealing approach to police district design. Comput Oper Res 29(6):667–684. https://doi.org/10.1016/s0305-0548(01)00056-9
Garey MR, Johnson DS (1977) The rectilinear Steiner tree problem is NP-complete. SIAM J Appl Math 32(4):826–834
Garey MR, Johnson DS (1979) Computers and intractability: a guide to the theory of NP-completeness. W. H. Freeman & Co., New York
Geofabrik GmbH, OpenStreetMap Contributors (2020) OpenStreetMap data Germany. https://download.geofabrik.de/europe/germany.html. Accessed 06 Aug 2020
Green L, Kolesar P (1989) Testing the validity of a queueing model of police patrol. Manage Sci 35(2):127–148. https://doi.org/10.1287/mnsc.35.2.127
Green LV, Kolesar PJ (2004) Anniversary article: improving emergency responsiveness with management science. Manage Sci 50(8):1001–1014. https://doi.org/10.1287/mnsc.1040.0253
Haase K, Müller S (2014) Upper and lower bounds for the sales force deployment problem with explicit contiguity constraints. Eur J Oper Res 237(2):677–689. https://doi.org/10.1016/j.ejor.2014.01.061
Hess SW, Weaver JB, Siegfeldt HJ et al (1965) Nonpartisan political redistricting by computer. Oper Res 13(6):998–1006. https://doi.org/10.1287/opre.13.6.998
Iain D, Joey H, Miles L (2017) Jump: a modeling language for mathematical optimization. SIAM Rev 59(2):295–320. https://doi.org/10.1137/15M1020575
Kalcsics J, Ríos-Mercado RZ (2019) Districting problems. In: Laporte G, Nickel S, Saldanha da Gama F (eds) Location science. Springer International Publishing, Cham, pp 705–743. https://doi.org/10.1007/978-3-030-32177-2_25
Kalcsics J, Nickel S, Schröder M (2005) Towards a unified territorial design approach—applications, algorithms and GIS integration. TOP 13(1):1–56. https://doi.org/10.1007/BF02578982
Kariv O, Hakimi SL (1979) An algorithmic approach to network location problems. II: the p-medians. SIAM J Appl Math 37(3):539–560. https://doi.org/10.2307/2100911
Kern GM (1989) A computer simulation model for the study of police patrol deployment. Simulation 52(6):226–232. https://doi.org/10.1177/003754978905200606
Kolesar P, Walker W (1975) A simulation model of police patrol operations: program description. RAND Corporation
Larson RC (1974) A hypercube queuing model for facility location and redistricting in urban emergency services. Comput Oper Res 1:67–95. https://doi.org/10.1016/0305-0548(74)90076-8
Li W, Goodchild MF, Church R (2013) An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems. Int J Geogr Inf Sci 27(6):1227–1250. https://doi.org/10.1080/13658816.2012.752093
Liberatore F, Camacho-Collados M (2016) A comparison of local search methods for the multicriteria police districting problem on graph. Math Probl Eng 2016:1–13. https://doi.org/10.1155/2016/3690474
Liberatore F, Camacho-Collados M, Vitoriano B (2020) Police districting problem: literature review and annotated bibliography. In: Ríos-Mercado RZ (ed) International series in operations research & management science: optimal districting and territory design. Springer International Publishing, Cham, pp 9–29. https://doi.org/10.1007/978-3-030-34312-5_2
Mayer A (2009) Geospatial technology helps east orange crack down on crime. Q Bull Appl Geogr Study Crime Public Saf 1(4):8–10

Mayorga ME, Bandara D, McLay LA (2013) Districting and dispatching policies for emergency medical service systems to improve patient survival. IIE Trans Healthc Syst Eng 3(1):39–56. https://doi.org/10.1080/19488300.2012.762437

McCabe JE (2017) Factors related to police staffing. Int J Humanit Soc Sci Rev 3(6):1–14

Mehrotra A, Johnson EL, Nemhauser GL (1998) An optimization based heuristic for political districting. Manage Sci 44(8):1100–1114. https://doi.org/10.1287/mnsc.44.8.1100

Miller HF, Knoppers BA (1972) Computer simulation of police dispatching and patrol functions. In: Cooper G (ed) International symposium on criminal justice information and statistics systems proceedings. National Institute of Justice, p 167–179

Mitchell PS (1972) Optimal selection of police patrol beats. J Crim Law Criminol Police Sci 63(4):577–584. https://doi.org/10.2307/1141814

Önal H, Wang Y, Dissanayake ST et al (2016) Optimal design of compact and functionally contiguous conservation management areas. Eur J Oper Res 251(3):957–968. https://doi.org/10.1016/j.ejor.2015.12.005

Peet R (1975) The geography of crime: a political critique. Prof Geogr 27(3):277–280. https://doi.org/10.1111/j.0033-0124.1975.00277.x

Reese J (2006) Solution methods for the p-median problem: an annotated bibliography. Networks 48(3):125–142. https://doi.org/10.1002/net.20128

Regis-Hernández F, Lanzarone E, Bélanger V et al (2018) A data-driven districting to improve emergency medical service systems. IFAC-PapersOnLine 51(11):998–1003. https://doi.org/10.1016/j.ifacol.2018.08.477

Ríos-Mercado RZ (2020) Research trends in optimization of districting systems. In: Ríos-Mercado RZ (ed) International series in operations research & management science: optimal districting and territory design. Springer International Publishing, Cham, pp 3–8. https://doi.org/10.1007/978-3-030-34312-5_1

Salazar-Aguilar MA, Ríos-Mercado RZ, Cabrera-Ríos M (2011) New models for commercial territory design. Netw Spat Econ 11(3):487–507. https://doi.org/10.1007/s11067-010-9151-6

Shen Y, Potvin JY, Rousseau JM et al (1995) A computer assistant for vehicle dispatching with learning capabilities. Ann Oper Res 61(1):189–211. https://doi.org/10.1007/BF02098288

Shirabe T (2009) Districting modeling with exact contiguity constraints. Environ Plann B Plann Des 36(6):1053–1066. https://doi.org/10.1068/b34104

Souza de RM, Morabito R, Chiyoshi FY et al (2015) Incorporating priorities for waiting customers in the hypercube queuing model with application to an emergency medical service system in brazil. Eur J Oper Res 242(1):274–285. https://doi.org/10.1016/j.ejor.2014.09.056

Swersey AJ (1994) The deployment of police, fire, and emergency medical units. In: Handbooks in operations research and management science: operations research and the public sector. Elsevier, pp. 151–200, https://doi.org/10.1016/S0927-0507(05)80087-8

van den Berg PL, Legemaate GAG, van der Mei RD (2017) Increasing the responsiveness of firefighter services by relocating base stations in Amsterdam. Interfaces 47(4):352–361. https://doi.org/10.1287/inte.2017.0897

Vestby A, Vestby J (2021) Machine learning and the police: asking the right questions. Polic J Policy Pract 15(1):44–58. https://doi.org/10.1093/police/paz035

Wang J, Kwan MP (2018) Hexagon-based adaptive crystal growth Voronoi diagrams based on weighted planes for service area delimitation. ISPRS Int J Geo Inf 7:257. https://doi.org/10.3390/ijgi7070257

Wang Y, Önal H, Fang Q (2018) How large spatially-explicit optimal reserve design models can we solve now? An exploration of current models' computational efficiency. Nat Conserv 27:17–34. https://doi.org/10.3897/natureconservation.27.21642

Yang B, Viswanathan K, Lertworawanich P et al (2004) Fire station districting using simulation: case study in Centre region, Pennsylvania. J Urban Plan Dev 130(3):117–124. https://doi.org/10.1061/(asce)0733-9488(2004)130:3(117)

Zhang Y, Brown D (2014) Simulation optimization of police patrol districting plans using response surfaces. Simulation 90(6):687–705. https://doi.org/10.1177/0037549714533159

Zhang Y, Brown DE (2013) Police patrol districting method and simulation evaluation using agent-based model & GIS. Secur Inform 2(1):55. https://doi.org/10.1186/2190-8532-2-7

Zhu S, Wang H, Xie Y (2021) Data-driven optimization for atlanta police zone design. https://doi.org/10.48550/ARXIV.2104.00535, arXiv:2104.00535

Zoltners AA, Sinha P (1983) Sales territory alignment: a review and model. Manage Sci 29(11):1237–1256. https://doi.org/10.1287/mnsc.29.11.1237

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.