

The Edge-based Contiguous p -median Problem with Connections to Logistics Districting *

Zeyad Kassem ^a and Adolfo Escobedo ^b

^a School of Computing and Augmented Intelligence, Arizona State University,
Tempe, Arizona, 85281

^b Edward P. Fitts Department of Industrial and Systems Engineering,
North Carolina State University, Raleigh, North Carolina, 27606

Abstract

This paper introduces the edge-based contiguous p -median (EC p M) problem to partition the roads in a network into a given number of compact and contiguous territories. Two binary programming models are introduced, both of which incorporate a network distance. The first model requires an exponential number of cut set-based constraints to model contiguity; it is paired with a separation scheme that usually generates only a small number of these constraints, namely, a branch-and-cut (B&C) algorithm. The second model utilizes a polynomial number of shortest-path constraints to model contiguity and can be solved with off-the-shelf solvers. The respective solution approaches are tested on road networks with over 2,700 nodes and close to 3,400 edges, yielding models with over 9.6 million binary variables. Solving the model based on shortest path contiguity (SPC) constraints via standard branch and bound attains speedups in computational time of up to 17x relative to the cut set-based B&C implementation. In addition, the SPC constraints are demonstrated to be supervalid inequalities of the edge-based p -median (E p M) model (i.e., for which contiguity is not explicitly required), meaning that they may cut off integer-feasible solutions and some, but not all, of the optimal solutions of this simpler problem. Finally, the paper explores structural insights and connections between EC p M and the edge-based districting (EBD) problem, which enforces an additional work balance criterion. An existing model that utilizes cut set-based contiguity constraints was unable to find a feasible solution within 12 hours for any of the tested instances, while an SPC-based EBD model was able to solve most of these to optimality.

Keywords: location analysis; last-mile logistics; p -median; districting; supervalid inequalities

1 Introduction

Over the last few years, there has been a global acceleration in the transition to the e-commerce business model. This has led to increased interest in last-mile freight logistics,

*A preliminary version of this paper appeared in conference proceedings in (Kassem and Escobedo, 2023a). The associated conference paper contained the two formulations for the edge-based contiguous p -median problem. All other content presented herein, including theorems, proofs, algorithms, computational results, insights, and writing, differs from that conference paper.

that is, the final leg of a delivery service in a consumer supply chain network where the consignment is sent directly to the consumer or to collection points from a transportation hub (Bosona, 2020). The last mile is the least efficient part of the supply chain, with costs contributing up to 28% of the total transportation cost (Bergmann et al., 2020). Moreover, last-mile logistics can contribute up to 30% of vehicle miles and up to 50% of greenhouse gas (GHG) emissions in developed urban areas (Dablanc and Rakotonarivo, 2010). Prior studies show that proximity from warehouses to service areas is a dominant factor in operational and environmental costs (Wygonik and Goodchild, 2018). In addition, the consolidation of shipments to customers along road segments and routes can help attain economies of density. This, in turn, enables the utilization of larger and more cost-effective vehicles, helping distribute fixed costs among more units and lowering variable costs, reducing the unit shipment cost (Alumur et al., 2021).

Discrete location-allocation models provide a systematic framework for prioritizing such key considerations in the planning of service areas onto which last-miles logistics are to be subsequently deployed. These models determine simultaneously how to locate a set of facilities and/or services (e.g., warehouses, schools, hospitals) and allocate customers or communities to the chosen locations according to one or several decision criteria (Marianov et al., 2002). They are usually defined on graphs, which can represent various types of territories — road networks, air transport networks, maritime networks, etc. (Tansel et al., 1983) — and have been adapted for different types of distance functions (Brimberg et al., 2008). Discrete location-allocation models that employ discrete (i.e., network) distances are especially suited to reflect practical aspects of road travel encountered in last-miles logistics (Kalcsics and Ríos-Mercado, 2019). A quintessential model with these characteristics is the p -median, which was first introduced in Hakimi (1964). The model selects p facilities from among a discrete set of candidate locations—most commonly represented as the nodes of the territory graph—and allocates customers to these facilities, with the objective of minimizing the sum of the weighted distances from each facility to its allocated customers (Hakimi, 1964).

Prior research has introduced numerous models and algorithms for the p -median, an NP-hard problem (Hakimi, 1979). Thorough reviews from the facility location literature are available in Owen and Daskin (1998), Melo et al. (2009), and Ulukan and Demircioğlu (2015); textbook treatments include Drezner and Hamacher (2004) and Daskin (2013). Similar to Hakimi (1964), many works employ network distances (or a function thereof) between the facilities and customers as a proxy for operational costs. To reflect various practical considerations, the p -median has been extended to impose additional restrictions. For instance, the capacitated p -median problem introduces limits on the amount of demand that each facility can serve (Ceselli and Righini, 2005). Another example is the dynamic p -median problem, which allows facilities to be opened, closed, relocated, or mobilized over a planning horizon, optimizing both location and timing decisions to minimize the total costs of satisfying the demand of allocated customers (Güden and Süral, 2019).

While various practical considerations have been incorporated into p -median models, other essential aspects have not been widely explored. The related literature tends to concentrate on node-based partitioning models, meaning that the customers (and facilities) are assumed to be situated on the vertices of the underlying territory graph (i.e., street intersections). Alternative partitioning models for reflecting the locations of residential customers on the graph’s edges (i.e., along the streets), which is relevant to various

logistics contexts (Kalcsics and Ríos-Mercado, 2019), have received significantly less attention. While the related edge-based graph partitioning problem has been recently studied in relation to computer science applications, e.g., distributed graph pre-processing on large graphs, characteristics of the models developed to address it are not tailored to logistics applications. Moreover, the computational techniques developed for the edge-based graph partitioning problem have focused on heuristic algorithms (Zhang et al., 2017; Mayer et al., 2018; Mayer and Jacobsen, 2021). Our work seeks to probe deeper into the study of edge-based location-allocation models and their solution techniques by considering the *edge-based p -median* problem and key extensions thereof. To this end, we introduce binary programming models and exact algorithms, and we analyze their theoretical and computational implications.

Another practical consideration that is rarely explicitly enforced in p -median models is contiguity, even though having a direct path from every facility to its allocated customers is inherently beneficial. An exception is the node-based p -median extension introduced by Lolonis and Armstrong (1993), which imposes this criterion via a set of non-linear expressions. The resulting model is incompatible with most optimization solvers; hence, the authors introduce a heuristic approach to solve it. Another exception is the connected p -median problem, which requires only the selected medians (i.e., facilities) to form a connected subgraph (Yen et al., 2010). Our work introduces different sets of linear constraints for imposing a more extensive type of contiguity that requires the selected center of each territory and its allocated edges to be connected. It is relevant to add that, although this type of contiguity is commonly enforced in other spatial optimization models (Williams, 2002; Carvajal et al., 2013; Salazar-Aguilar et al., 2011; García-Ayala et al., 2016), it has not been explicitly incorporated in more elementary location-allocation models. Addressing this research gap is important because certain design criteria associated with these more restrictive models can conflict with the basic requirements of p -median models. For example, the requirement that each facility must serve approximately the same amount of customers or demand, which is commonplace in logistics districting (see below), can degrade compactness of the service areas (García-Ayala et al., 2016).

Certain extensions of the p -median problem implicitly yield contiguous territories, but they come with excessive computational costs relative to planning purposes. In particular, the Hamiltonian p -median (HpM) problem is a location-routing problem that combines the p -median problem and the traveling salesperson problem (Branco and Coelho, 1990). HpM simultaneously selects p depots, allocates customers to one of the depots, and determines routes for serving customers allocated to the same depot, with the objective of minimizing total distance traveled. It is related to but more challenging than the generalized vehicle routing problem, in which the depots are predetermined (Gollowitzer et al., 2011). The inclusion of subtour elimination constraints to determine each route in HpM effectively imposes contiguity, by enforcing each vehicle to complete a tour that visits each of its assigned customers exactly once during a single trip that starts and ends at the depot. Hence, it is a more restrictive form of contiguity than requiring simply the customers served by each depot to be connected. Logistics districting represents an important class of location-allocation models that also build on the p -median, imposing the added requirements of contiguity and work balance, and exacerbate its computational difficulties. Before proceeding, it is important to distinguish these models from those developed in political districting, whose study has received comparatively more attention spanning several decades (e.g., Hess et al.

(1965); Mehrotra et al. (1998); Ricca et al. (2011); Validi et al. (2022)). While their basic design criteria—compactness, contiguity, and work balance (see Section 2)—are conceptually similar, the two types of models exhibit substantive differences including the basic units to be partitioned, distance functions employed, and additional design criteria that are unique to each context.

This paper makes several contributions to territorial design. First, it introduces the edge-based contiguous p -median (EC p M) problem and two exact binary formulations to solve it. The formulations partition a road network into a fixed number of compact and contiguous territories by locating centers and allocating roads to these centers, with the objective of minimizing the total distance within each territory. The two proposed formulations differ in how they enforce contiguity: the first requires an exponential number of cut set constraints, while the second utilizes a polynomial number of shortest-path constraints. As a second contribution, this paper introduces a separation algorithm that generates only a small number of cut set constraints to solve the first model, namely, a branch-and-cut (B&C) algorithm. The second model is solved using off-the-shelf methods (i.e., branch-and-bound). As a third contribution, the paper derives three logically equivalent sets of constraints for enforcing shortest-path contiguity (SPC) and compares them using polyhedral techniques. The SPC constraints are also demonstrated to be *supervalid inequalities* for the simpler edge-based p -median problem (E p M), meaning that they may cut off integer-feasible solutions and some, but not all, of the optimal solutions for this simpler problem. As a fourth contribution, the paper carries out experiments to test the computational impacts of applying the SPC-based model on road networks with over 2,700 nodes and close to 3,400 edges. As a final contribution, the paper explores theoretical and computational implications of the proposed p -median models on edge-based districting (EBD).

The remainder of the paper is structured as follows. Section 2 provides the background for this paper and a brief overview of related works. Section 3 introduces the E p M and EC p M problems and binary programming formulations to solve them. It also derives additional theoretical insights regarding the SPC constraints introduced for EC p M and performs computational studies to test the featured models and solution schemes. Section 4 explores the connections between EC p M and EBD, and it tests the additional computational benefits of the proposed SPC-based models. Finally, Section 5 concludes the paper.

2 Background

Location-allocation models are designed to create territories that satisfy desirable criteria. A territory is considered to be *compact* if its shape is nearly circular or approximately square-shaped, non-distorted, without holes, and with smooth boundaries (Butsch et al., 2014). By prioritizing compactness, location-allocation models can help improve the subsequent road-based routing for e-commerce companies, as compact territories lead to shorter travel distances (García-Ayala et al., 2016). Another design criterion that can improve last-mile logistics is contiguity. A territory is considered to be *contiguous* if it is possible to travel between any two points within a territory without leaving the territory. Contiguity can enable e-commerce companies to aggregate different customer orders placed in close proximity to each other and eventually deploy their routing algorithms within each non-overlapping territory (Álvarez-Miranda and Pereira, 2021). Contiguity has been studied in spatial optimization, and it is among the most difficult criteria to enforce in these con-

texts (Ricca and Simeone, 2008). This criterion can be implicitly prioritized by optimizing compactness; however, compact territories need not be contiguous (Shirabe, 2009).

There have been several attempts to impose contiguity across different applications including political districting (Garfinkel and Nemhauser, 1970; Validi et al., 2022), land acquisition (Williams, 2002), forest planning (Carvajal et al., 2013), and territorial districting (Salazar-Aguilar et al., 2011; García-Ayala et al., 2016). A natural approach is to leverage similar restrictions from other optimization models such as subtour elimination constraints from the traveling salesperson problem (TSP) (Dantzig et al., 1954)—which prevent solutions that induce cycles with fewer arcs than the total number of cities to be visited. Subtour elimination constraints can be imposed through *cut set* logic, which requires that a feasible solution must include at least one arc between every strict subset of cities and its complement. Variations of this concept have been devised for node-based (Salazar-Aguilar et al., 2011) and edge-based (García-Ayala et al., 2016) districting models. However, the inclusion of the full constraint set, which has exponential cardinality, is computationally prohibitive except for very small instances. Thus, these models cannot be implemented with off-the-shelf methods (i.e., branch-and-bound) and, consequently, they tend to be paired with a separation scheme.

There are three main types of separation schemes for dealing with contiguity: iterative branch-and-bound with a cut generation scheme (B&B&Cut), integer separation, and fractional separation. B&B&Cut is an iterative exact algorithm that solves a relaxed version of the original model without contiguity constraints using branch-and-bound (B&B) (Salazar-Aguilar et al., 2011). Then, it uses a search algorithm—e.g., breadth first search or depth first search—to check for disconnected territories, adds the needed cuts, and re-solves the new model using B&B. At each iteration, the search algorithm either finds some violated contiguity constraints, which are added to the model, or it concludes that all territories are contiguous, terminating the solution algorithm. On the other hand, integer separation solves the relaxed version of the original model without the contiguity constraints, and then it deploys a cutting plane algorithm within the B&B exploration for every incumbent integer solution. That is, once an integer solution is obtained, the algorithm checks if there are violated contiguity constraints using a search algorithm, and it adds the needed cuts on-the-go (e.g., as lazy constraints) without having to restart B&B every time a violated contiguity constraint is found (Mendes et al., 2022). The implementation of integer separation requires more sophisticated knowledge of the optimization solver, for example, to issue callbacks for adding the violated constraints. Finally, fractional separation deploys a cutting plane algorithm within B&B exploration for every fractional solution. This strategy relies on complex and/or computationally expensive algorithms (e.g., multiple calls to the minimum cut algorithm) (Validi et al., 2022). Therefore, prior works have implemented primarily B&B&Cut or integer separation (Salazar-Aguilar et al., 2011; García-Ayala et al., 2016; Mendes et al., 2022; Validi et al., 2022). Branch-and-cut algorithms tend to align with the second and third types of separation scheme (Martin, 2001).

Previous research has also introduced compact constraints sets (i.e., of polynomial cardinality) for imposing contiguity. These are divided into two main types: network flow-based (Shirabe, 2009; Jafari and Hearne, 2013; Wang et al., 2021) and neighbor-based (Mehrotra et al., 1998; Önal et al., 2016; Farughi et al., 2020). With the former type, the undirected planar graph is converted into a bidirected graph by replacing each undirected edge with two arcs (one for each direction). Contiguity is achieved by ensuring flow between nodes

occupying the same territory. Models equipped with these constraints tend to perform poorly computationally, except on small problem instances (Wang et al., 2021; Validi et al., 2022). Within each territory, different network flow configurations can represent the same allocation solution, causing a high level of symmetry. Thus, these models are usually augmented with symmetry-reducing cutting planes (Wang et al., 2021; Mendes et al., 2022). An alternative type of compact contiguity constraints considers logic on neighbors, specifically, if a node is allocated to a certain center node, then at least one of the neighboring nodes must be allocated to the same center node. Propagating this requirement generates a chain of mutually adjacent nodes, forming a path between the center node and any node assigned to it within each territory. Such constraints can be regarded as a special application of dominance constraints developed for the simple plant location problem (Cánovas et al., 2007). To the best of our knowledge, the comparative performance of these compact contiguity constraints vis-à-vis cut set-based constraints has not been directly evaluated.

3 Exact Formulations

In this section, we define the $ECpM$ problem and present a binary programming model with cut set-based contiguity constraints for solving it. The latter constraints are exponential in number, and their full specification requires prohibitive computational resources. Therefore, we complement this model with a branch-and-cut (B&C) algorithm that tends to generate only a small subset of these constraints on real-world instances. We also propose an alternative model that incorporates shortest-path contiguity (SPC) constraints, which are polynomial in number. In addition, we demonstrate that the SPC constraints represent supervalid inequalities of EpM (i.e., $ECpM$ without the contiguity requirement).

3.1 The Edge-based p -median (EpM) Model

Prior to introducing the $ECpM$ problem, we introduce the EpM problem, which aims to determine the optimal location of p center nodes and the edges that are allocated to them. The underlying road network is represented as an undirected planar graph $G = (V, E)$, which is assumed to be connected. To proceed with the mathematical formulation of EpM , let $l_{(m,n)}$ be the length of edge $(m,n) \in E$ and consider the following network distance concepts.

Definition 3.1. Let $SP_{i,j}$ be the set of edges that make up the shortest path from node i to node j . This parameter is expressed mathematically as

$$SP_{i,j} = \arg \min_{\rho \in \mathcal{P}_{i,j}} \left\{ \sum_{(m,n) \in \rho} l_{(m,n)} \right\},$$

where $\mathcal{P}_{i,j}$ is the set of all possible paths between nodes $i, j \in V$.

Definition 3.2. Let $SP_{i,(j,k)}$ be the set of edges that make up the shortest path from node i to edge (j,k) ; it is given by either $SP_{i,j}$ or $SP_{i,k}$, depending on which of j or k is the tail node in the path. This parameter is expressed mathematically as

$$SP_{i,(j,k)} = \arg \min \left\{ \sum_{(m,n) \in SP_{i,j}} l_{(m,n)}, \sum_{(m,n) \in SP_{i,k}} l_{(m,n)} \right\}.$$

Definition 3.3. Let $d_{i,(j,k)}$ be a network distance equal to the sum of positive edge lengths that make up the shortest path from node i to edge (j, k) . This parameter is expressed mathematically as

$$d_{i,(j,k)} = \sum_{(m,n) \in SP_{i,(j,k)}} l_{(m,n)},$$

The shortest path between every pair of nodes in G can be computed efficiently using Dijkstra's algorithm (Ahuja et al., 1988). Based on Definitions 3.1 and 3.2, this result can be extended to the shortest path between every node and edge in the graph, which must be computed to construct the EpM model. The components of the model are described in the ensuing paragraphs.

Sets and Indices:

$i \in V$ Nodes in the network (corresponding to street crossings, dead ends, etc.).

$(j, k) \in E$ Edges in the network (corresponding to roads, streets, etc.).

Parameters:

$d_{i,(j,k)}$ Distance from node i to edge (j, k) (see Definition 3.3).

p Number of territories.

Decision Variables:

$$w_i = \begin{cases} 1 & \text{if node } i \text{ is selected as a center node,} \\ 0 & \text{otherwise.} \end{cases}$$

$$x_{i,(j,k)} = \begin{cases} 1 & \text{if edge } (j, k) \text{ is allocated to center node } i, \\ 0 & \text{otherwise.} \end{cases}$$

$$(EPM) \quad \min \sum_{i \in V} \sum_{(j,k) \in E} d_{i,(j,k)} x_{i,(j,k)} \quad (1)$$

$$\text{s.t.} \quad \sum_{i \in V} x_{i,(j,k)} = 1 \quad \forall (j, k) \in E \quad (2)$$

$$\sum_{i \in V} w_i = p \quad (3)$$

$$x_{i,(j,k)} \leq w_i \quad \forall i \in V, \forall (j, k) \in E \quad (4)$$

$$x_{i,(j,k)} \in \mathbb{B} \quad \forall i \in V, \forall (j, k) \in E \quad (5)$$

$$w_i \in \mathbb{B} \quad \forall i \in V \quad (6)$$

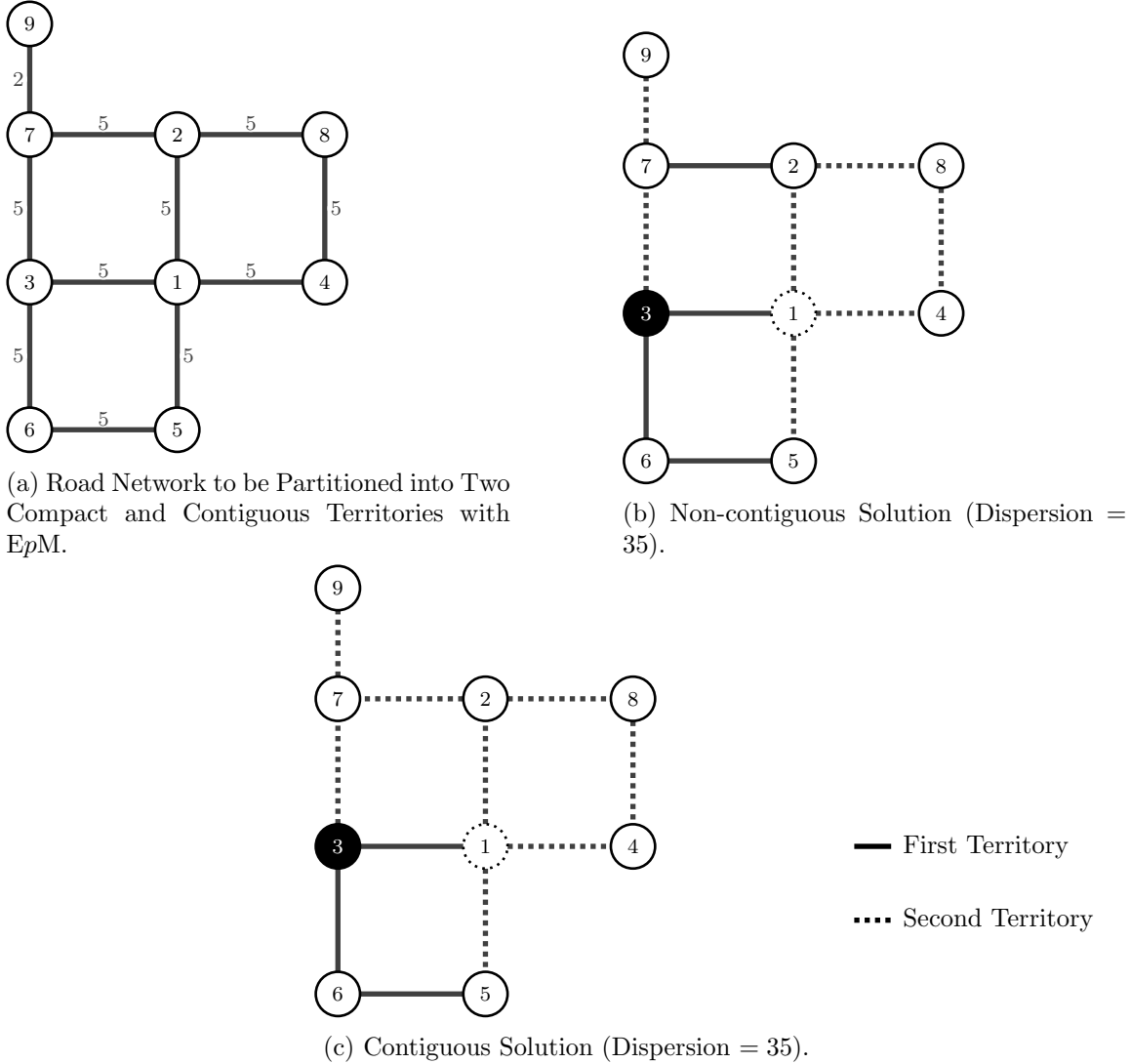


Figure 1: Road Network and E_pM Solutions with Identical Dispersions (i.e., Objective Function Values)

Objective Function (1) maximizes compactness by minimizing the dispersion from each center node i to its allocated edges. Constraint (2) ensures that each edge $(j, k) \in E$ is allocated to exactly one center node. Constraint (3) ensures that exactly p center nodes are selected. Constraint (4) allows edge (j, k) to be allocated to node i only if it is selected to be a center node. Constraints (5) and (6) specify the domain of the variables.

Next, Example 3.1 and Figure 1 illustrate that model (1)-(6) can return solutions corresponding to non-contiguous territories. Then, the ensuing subsection presents two binary programming models for extending E_pM to exclude such outcomes; the models differ in how they impose contiguity.

Example 3.1. Consider the connected planar graph $G = (V, E)$ depicted in Figure 1a, where $|V| = 9$, $|E| = 11$, and $p = 2$. Figures 1b and Figure 1c display two E_pM solutions. In each case, solid edges correspond to the first territory, whereas dotted edges correspond

to the second territory; the centers of the two territories follow the same formatting scheme as the edges. In both of solutions, nodes 1 and 3 are the centers of the first and second district, respectively. The sole difference is the allocation of edge (2, 7), which is allocated to the first territory in Figure 1b and to the second territory in Figure 1c. The two solutions have the same dispersion (35), but only the latter is contiguous.

3.2 The Edge-based Contiguous p -median Model with Cut Set Contiguity Constraints (EC p M-CSC)

As an extension of the E_p M problem, this subsection presents a binary programming formulation to partition E into p compact and contiguous territories. The formulation imposes contiguity by forcing every cut set that separates any pair of nonadjacent edges allocated to a certain territory to allocate at least one edge from the cut set to the same territory. Cut set contiguity constraints, which are exponential in number, are analogous to subtour elimination constraints of the TSP (see Section 2). They were first devised for node-based districting by Drexler and Haase (1999), who also explain that it is usually possible to solve real-world instances by imposing only a small number of these constraints. Cut set contiguity constraints are adapted for edge-based districting in García-Ayala et al. (2016). Note, however, that their model assumes that the centers are predetermined (i.e., it is an allocation-only model).

Sets and Indices:

$S \subset E$ A subset of edges.

$V(S) \subset V$ Set of nodes that are incident to any edge in S .

$\sigma(S)$ Cut set of S (i.e., $\sigma(S) := \{(i, j) \in E \mid i \in V(S), j \in V \setminus V(S)\}$).

$$\text{(EC}p\text{M-CSC)} \quad \min \sum_{i \in V} \sum_{(j,k) \in E} d_{i,(j,k)} x_{i,(j,k)} \quad (7)$$

$$\text{s.t.} \quad (2) - (6)$$

$$\sum_{s \in \sigma(S)} x_{is} - \sum_{s \in S} x_{is} \geq x_{i,(j,k)} - |S| \quad \forall i \in V, \forall (j,k) \in E, S \subset E \setminus \sigma(\{(j,k)\}) \quad (8)$$

The main difference from E_p M is the addition of Constraints (8), which ensure that any edge (j, k) allocated to center node i should be adjacent to other edges allocated to the same center node. To elaborate, let $S \subset E \setminus \sigma(\{(j, k)\})$ be any subset of edges not adjacent to (j, k) , where $\sigma(\{(j, k)\})$ is the set of adjacent edges to (j, k) . If edge (j, k) is not allocated to center node i (that is, $x_{i,(j,k)} = 0$), the constraint is always satisfied. Additionally, if there is one or more edges in S not allocated to center node i , then the value of second term is strictly less than $|S|$, and the constraint is also always satisfied. The constraint is activated only when all edges in S are allocated to center node i and edge (j, k) is allocated to center node i ($x_{i,(j,k)} = 1$). It imposes this logic by requiring that at least one edge in the cut set of set S must be allocated to center node i . Applying these constraints for every possible subset S ensures that all edges within each service area are connected.

3.2.1 Branch-and-Cut Scheme

There is an exponential number of Constraints (8), making their full implementation impractical. Next, we pair EC p M-CSC with an integer separation scheme, namely a branch-and-cut algorithm (B&C), which is adapted from an algorithm introduced in Mendes et al. (2022). B&C generates only those cut set contiguity constraints that are deemed necessary based on the solution to a relaxed version of EC p M. These constraints are determined based on the output of a separation algorithm, whose pseudocode is given in Algorithm 1. Note that, in the B&C implementation, the cuts corresponding to (8) are added on-the-go as *lazy constraints*, thereby avoiding having to restart B&B to remove point \mathbf{x}^* from the current relaxation.

To describe Algorithm 1, let $G[E_k] = (V_k, E_k)$ denote the subgraph of G induced from the edges in territory k , denoted as E_k , which gives that $V_k := \{u, v \in V \mid (u, v) \in E_k\}$. The algorithm first solves a relaxed version of the original model omitting all contiguity constraints. In the next and all subsequent iterations, the algorithm solves the current relaxation of EC p M-CSC, yielding an incumbent binary solution \mathbf{x}^* . The solution is inspected for violations in contiguity using breadth first search (BFS) (or another comparable search algorithm), which returns the connected components in each service area. When there is more than one connected component within an area, a violation in contiguity is detected, and the algorithm adds a set R of cut set constraints to remove it. Referring to the pseudocode, Lines 2-11 iterate over each of the p territories found from the relaxation solution. When territory k is not contiguous, Lines 5-11 add the violated contiguity constraints for every pair S'_k, S''_k of connected components in E_k . Specifically, Line 9 adds one violated contiguity constraint, for each $(l, m) \in S'_k$ that is found to be disconnected from the edges within the connected component S''_k ; the added constraint enforces that at least one edge in the cut set of S''_k must be allocated to the same territory as edge (l, m) . Lines 10-11 carry out the same function as Lines 8-9, but for each $(u, v) \in S''_k$.

3.3 The Edge-based Contiguous p -median Model with Shortest-path Contiguity Constraints (EC p M-SPC)

An alternative approach to model contiguity used in spatial optimization applications is to leverage shortest paths. Zoltners and Sinha (1983) introduce the concept of hierarchical adjacency trees for the sales territory alignment problem with predetermined centers. For each predetermined center node, a set of shortest paths is determined from the center to all other nodes, with the predecessor-successor along each shortest path recorded. Contiguity is enforced by ensuring that no node can be allocated to a center node unless one or more of its immediate preceding nodes along a shortest path is also allocated to the same center node. Cova and Church (2000) derive shortest-path contiguity- n constraints for single-region site search problems. These constraints allow each spatial unit to be reached from the root along only one of its n -shortest paths; to that end, it is necessary to compute the shortest, the second-shortest, and so on, up to the n^{th} -shortest path from each spatial unit to the root, where n is assumed to be reasonably small. Mehrotra et al. (1998) develop and implement a special version of shortest-path contiguity- n constraints for political districting where only the first shortest path is considered. More recently, Farughi et al. (2020) introduce a distinct set of shortest-path contiguity constraints for designing health care districts. The

Algorithm 1 Separation Algorithm

Input:
 \mathbf{x}^* , the binary solution obtained from the current relaxation of ECpM-CSC

 $\{E_1, E_2, \dots, E_p\}$, a p -partition of E
 $\sigma(\{(u, v)\})$, $\forall (u, v) \in E$, the cut set of each edge

Output: a set R , which is empty if all p territories are contiguous; otherwise it is the set of valid inequalities that cuts off \mathbf{x}^* from the current relaxation space

```

1:  $R \leftarrow \emptyset$ 
2: for  $k = 1$  to  $p$  do
3:   Deploy BFS within  $G[E_k]$  to identify whether  $G[E_k]$  is contiguous
4:   Identify all of the connected components, say,  $S_k^1, S_k^2, \dots, S_k^t$ 
5:   if  $G[E_k]$  is not contiguous then
6:     Identify center for district  $k$ , and denote its index as  $i$ 
7:     for all pairs  $S'_k, S''_k$  of connected components within  $E_k$  do
8:       for each  $(l, m) \in S'_k$ 
9:          $R \leftarrow R \cup \{\sum_{s \in \sigma(S''_k)} x_{is} - \sum_{s \in S''_k} x_{is} \geq x_{i,(l,m)} - |S''_k|\}$ 
10:      for each  $(u, v) \in S''_k$ 
11:         $R \leftarrow R \cup \{\sum_{s \in \sigma(S'_k)} x_{is} - \sum_{s \in S'_k} x_{is} \geq x_{i,(u,v)} - |S'_k|\}$ 
12:   end for
13: return  $R$ 

```

constraints ensure that if nodes j and k are allocated to center node i , then all nodes along $SP_{j,k}$ must also be assigned to center node i . A common feature of these shortest-path constraints is that they are defined exclusively for node-based territorial planning models. An exception is Cova and Church (2000), where the spatial unit is a raster, which nevertheless can be readily represented as a node.

To the best of our knowledge, shortest-path contiguity constraints—henceforth referred to as SPC constraints, for short—have not been defined for edge-based location-allocation models. This work seeks to fill this research gap. The main intuition behind the proposed constraints is that, for any connected planar graph with more than one edge, if edge (j, k) is assigned to center node i , then all edges along the shortest path that joins them must also be assigned to i . Since $SP_{i,(j,k)}$ is an uninterrupted path, requiring all of its edges to be allocated to center node i when edge (j, k) is allocated to i is tantamount to enforcing contiguity within each territory. It is possible to model this requirement in at least three logically equivalent ways, denoted as constraints SPC-1, SPC-2, and SPC-3; their expressions are given by

$$(SPC-1) |SP_{i,(j,k)}| x_{i,(j,k)} \leq \sum_{(l,m) \in SP_{i,(j,k)}} x_{i,(l,m)} \quad \forall i \in V, \forall (j, k) \in E \quad (10')$$

$$(SPC-2) x_{i,(j,k)} \leq x_{i,(l,m)} \quad \forall i \in V, \forall (j, k) \in E, \forall (l, m) \in SP_{i,(j,k)} \quad (10'')$$

$$(SPC-3) x_{i,(j,k)} \leq x_{i,(l,m)} \quad \forall i \in V, \forall (j, k) \in E, (l, m) \in SP_{i,(j,k)} \cap \sigma(\{(j, k)\}) \quad (10''')$$

where $\sigma(\{(j, k)\})$ is the cut set of edge (j, k) (i.e., the set of its adjacent edges). Note that in (10'''), (l, m) is the preceding edge to (j, k) on the shortest path from (j, k) to node i , because it belongs to both $SP_{i,(j,k)}$ and $\sigma(\{(j, k)\})$. Next, we establish that constraints sets

SPC-1, SPC-2, and SPC-3 are logically equivalent, and afterward we demonstrate which of the constraint sets induces the tightest formulation.

Theorem 3.1. For any instance of ECpM, constraints SPC-1, SPC-2, and SPC-3 are logically equivalent.

Proof. Proof. If $x_{i,(j,k)} = 0$, all three left-hand sides become 0, making the constraints redundant. If $x_{i,(j,k)} = 1$, the shortest path allocation requirement is imposed by each constraint set. SPC-1 forces the full set of edges along $SP_{i,(j,k)}$ to be allocated to node i , using a single inequality. SPC-2 relates, as separate expressions, the allocation to center i of each $(l, m) \in SP_{i,(j,k)}$ with the corresponding allocation variable $x_{i,(j,k)}$. SPC-3 follows a similar approach as SPC-2, but it enforces only that the edge that precedes $(j, k) \in SP_{i,(j,k)}$ must be allocated to i .

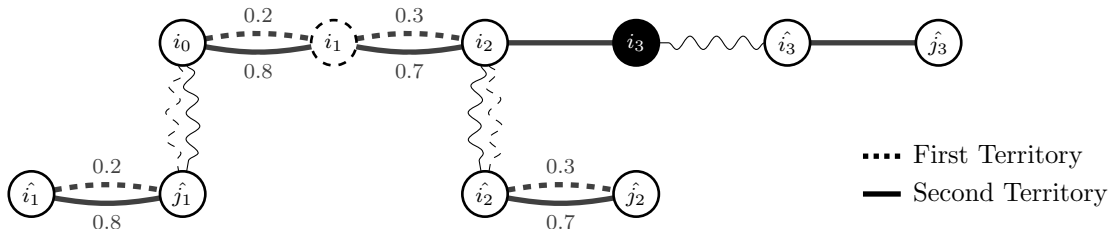


Figure 2: Subgraph Depicting Fractional Solutions for the Linear Relaxation of SPC-1 and SPC-2.

Theorem 3.2. Define the polytope $\mathcal{P}_{\text{SPC-1}} = \{(\mathbf{x}, \mathbf{w}) \in [0, 1]^{|V| \times |E|} \times [0, 1]^{|V|} : (\mathbf{x}, \mathbf{w}) \text{ satisfies (2)-(4), (10')}\}$ and the polytope $\mathcal{P}_{\text{SPC-2}} = \{(\mathbf{x}, \mathbf{w}) \in [0, 1]^{|V| \times |E|} \times [0, 1]^{|V|} : (\mathbf{x}, \mathbf{w}) \text{ satisfies (2)-(4), (10'')}\}$. For any instance of ECpM, $\mathcal{P}_{\text{SPC-2}} \subseteq \mathcal{P}_{\text{SPC-1}}$, and this inclusion can be strict.

Proof. Proof. Constraint (10') can be obtained as a linear combination of Constraint (10''); specifically, for any $i \in V$ and $(j, k) \in E$, summing (10'') over all $(l, m) \in SP_{i,(j,k)}$ yields Constraint (10'). This implies that all feasible solutions to $\mathcal{P}_{\text{SPC-2}}$ are also feasible to $\mathcal{P}_{\text{SPC-1}}$ and, thus, $\mathcal{P}_{\text{SPC-2}} \subseteq \mathcal{P}_{\text{SPC-1}}$.

We show that the inclusion can be strict. Let $p \geq 2$ and assume that there are at least two nodes with a shortest path of three hops (i.e., there are three consecutive edges separating them). We construct a solution $\mathbf{x}' \in \mathcal{P}_{\text{SPC-1}} \setminus \mathcal{P}_{\text{SPC-2}}$. An example is depicted in Figure 2 for convenience. Therein, squiggly lines indicate a collapsed non-empty set of edges along some indicated shortest path. A number adjacent to a dotted line indicates the value of the allocation variable associated with that edge and center node i_1 , and a number adjacent to a solid line indicates the value of the allocation variable associated with that edge and center node i_3 .

Without loss of generality, consider the shortest path from i_0 to i_3 , with $|SP_{i_0,i_3}| = 3$, and fix i_1 and i_3 as center nodes ($w_{i_3} = w_{i_1} = 1$). If $p > 2$, select the remaining $(p - 2)$ centers arbitrarily from $V \setminus \{i_1, i_3\}$; then set the allocation variables corresponding to edges (i_0, i_1) , (i_1, i_2) , and (i_2, i_3) as follows:

- Edge (i_0, i_1) . Set $x_{i_1,(i_0,i_1)} = 0.2$ and $x_{i_3,(i_0,i_1)} = 0.8$.
- Edge (i_1, i_2) . Set $x_{i_1,(i_1,i_2)} = 0.3$ and $x_{i_3,(i_1,i_2)} = 0.7$.

- Edge (i_2, i_3) . Set $x_{i_3, (i_2, i_3)} = 1$.

With consideration of the above three edges, proceed according to three mutually exclusive cases.

Case 1: $(i_0, i_1) \in SP_{i_1, (\hat{i}_1, \hat{j}_1)}$.

Set $x_{i_1, (\hat{i}_1, \hat{j}_1)} = 0.2$ and $x_{i_1, (l, m)} = 0.2 \forall (l, m) \in SP_{i_0, (\hat{i}_1, \hat{j}_1)}$; additionally, set $x_{i_3, (\hat{i}_1, \hat{j}_1)} = 0.8$, and $x_{i_3, (l, m)} = 0.8 \forall (l, m) \in SP_{i_0, (\hat{i}_1, \hat{j}_1)}$.

Case 2: $(i_1, i_2) \in SP_{i_1, (\hat{i}_2, \hat{j}_2)}$ and $(i_2, i_3) \notin SP_{i_1, (\hat{i}_2, \hat{j}_2)}$

Set $x_{i_1, (\hat{i}_2, \hat{j}_2)} = 0.3$ and $x_{i_1, (l, m)} = 0.3 \forall (l, m) \in SP_{i_2, (\hat{i}_2, \hat{j}_2)}$; additionally, set $x_{i_3, (\hat{i}_2, \hat{j}_2)} = 0.7$ and $x_{i_3, (l, m)} = 0.7 \forall (l, m) \in SP_{i_2, (\hat{i}_2, \hat{j}_2)}$.

Case 3: $(i_1, i_2), (i_2, i_3) \in SP_{i_1, (\hat{i}_3, \hat{j}_3)}$

Set $x_{i_3, (\hat{i}_3, \hat{j}_3)} = 1$ and $x_{i_3, (l, m)} = 1 \forall (l, m) \in SP_{i_3, (\hat{i}_3, \hat{j}_3)}$.

To finish constructing the solution, allocate all other edges in full to center i_1 .

The constructed solution \mathbf{x}' satisfies Constraints (2)-(4) and (10'): The sum of the allocation variables for each edge is equal to 1; the number of territories is equal to p ; and each edge is allocated to a node, only if the node is chosen as a center node.

Now, isolating Constraint (10') for edge (i_0, i_1) and node i_3 , the chosen variable settings give that

$$|SP_{i_3, (i_0, i_1)}| x_{i_3, (i_0, i_1)} = 2 \cdot (0.8) \leq 1 + 0.7 = x_{i_3, (i_1, i_2)} + x_{i_3, (i_2, i_3)}. \quad (9)$$

and, thus, the constraint is satisfied. On the other hand, Constraint (10'') is violated since

$$x_{i_3, (i_0, i_1)} = 0.8 \not\leq 0.7 = x_{i_3, (i_1, i_2)}. \quad (10)$$

□

Theorem 3.3. Let $\mathcal{P}_{\text{SPC-2}}$ be the polytope defined in Theorem 3.2 and define the polytope $\mathcal{P}_{\text{SPC-3}} = \{(\mathbf{x}, \mathbf{w}) \in [0, 1]^{|V| \times |E|} \times [0, 1]^{|V|} : (\mathbf{x}, \mathbf{w}) \text{ satisfies (2)-(4), (10''')}\}$. Then, $\mathcal{P}_{\text{SPC-2}} = \mathcal{P}_{\text{SPC-3}}$.

Proof. Proof. Constraints (10''') enforce that, if edge (j, k) is assigned to center node i , then the preceding edge (k, l) in $SP_{i, (j, k)}$ must also be assigned to i . By the same logic, if (k, l) is assigned to i , then the preceding edge (l, o) on the shortest path $SP_{i, (k, l)}$ must also be assigned to i . By transitivity, if $x_{i, (j, k)} \leq x_{i, (k, l)}$ and $x_{i, (k, l)} \leq x_{i, (l, o)}$, then $x_{i, (j, k)} \leq x_{i, (l, o)}$. The same logic can be extended to all adjacent pairwise edges along $SP_{i, (j, k)}$. Constraints (10'') enforce exactly the same logic, but they list all inequalities explicitly (i.e., without taking advantage of the implied transitivity). This means that $\mathcal{P}_{\text{SPC-2}}$ and $\mathcal{P}_{\text{SPC-3}}$ are equal. □

Upon inspection, there are $\mathcal{O}(|V|^2|E|)$ constraints of type SPC-2, and there are $\mathcal{O}(|V||E|)$ constraints of types SPC-1 and SPC-3. Because constraints SPC-3 induce a linear programming relaxation that is at least as tight as that obtained with constraints SPC-1 (see Theorem 3.2), the former are henceforth adopted for the remainder of this work. Therefore, the ECpM-SPC formulation is given by:

$$\text{(ECpM-SPC)} \quad \min \sum_{i \in V} \sum_{(j,k) \in E} d_{i,(j,k)} x_{i,(j,k)} \quad (11)$$

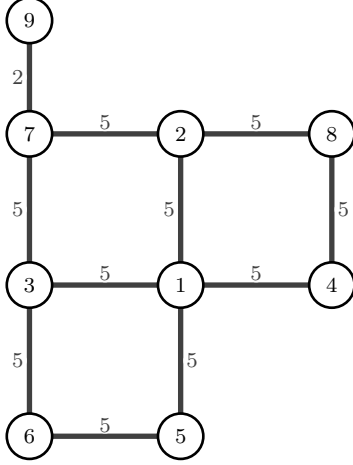
$$\begin{aligned} \text{s.t.} \quad & (2) - (6) \\ & x_{i,(j,k)} \leq x_{i,(l,m)} \quad \forall i \in V, \forall (j,k) \in E, (l,m) \in SP_{i,(j,k)} \cap \sigma(\{(j,k)\}) \end{aligned} \quad (12)$$

The contiguity criterion is not required by the Epm problem. Thus, at first glance, the SPC constraints may seem extraneous to this simpler model. Nevertheless, we show that they can also be useful for solving this problem. To that end, it is useful to differentiate between three types of cutting plane techniques that can be imposed to reduce the solution space. The best known and most widely applied type, *valid inequalities* (referred alternatively as valid logic cuts), can remove only fractional points of the relaxed feasible region, meaning that they do not remove any integer feasible points. On the other hand, non-valid logic cuts can make deeper cuts to the solution space by additionally allowing the removal of integer-feasible, non-optimal solutions. This type of cutting planes was first introduced by Hooker (1994) and have been applied in capacitated warehouse location (Osorio Lama and Mújica García, 1999), design of chemical processing networks (Hooker, 1994), and multilevel generalized assignment (Osorio and Laguna, 2003). The third type of cutting plane technique, known as supervalid inequalities, also allows the removal of non-optimal solutions, that are integer feasible (Israeli and Wood, 2002). However, they go one step further by allowing the removal of some, but not all, of the optimal solutions. In other words, they are a stronger type of cutting plane techniques than both valid inequalities and non-valid logic cuts. The latter type of cutting planes have been applied to interdiction problems (Seifair et al., 2017; Baycik and Sullivan, 2019; Wei and Walteros, 2024) and routing problems (Yuan et al., 2020, 2021).

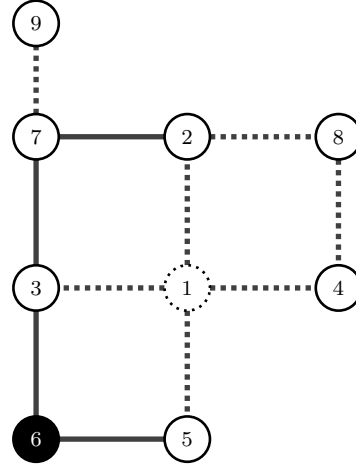
Theorem 3.4 shows that the SPC constraints are supervalid inequalities of Epm . Beforehand, the ensuing example motivates the featured insight regarding this additional use of the SPC constrains (see Figure 3).

Example 3.2. Consider the connected planar graph $G = (V, E)$ depicted in Figure 3a, where $|V| = 9$, $|E| = 11$, and $p = 2$. Figure 3b depicts a non-contiguous Epm solution, and Figure 3c shows an Epm solution satisfying the SPC constraints. In each case, solid edges correspond to the first territory, whereas dotted edges correspond to the second territory; the centers of the two territories follow the same formatting scheme as the edges. In both solutions depicted, nodes 1 and 6 are the centers. The sole difference is the allocation of edge $(2, 7)$: In 3b, it is allocated to the first territory, but in 3c it is allocated to the second territory. The latter allocation dominates the former, because $d_{1,(2,7)} = 5 < 10 = d_{6,(2,7)}$. The difference stems from the fact that allocating edge $(7, 9)$ to node 1 and enforcing SPC forces all edges along $SP_{1,(7,9)}$, including $(2, 7)$, to be allocated to center node 1. The following theorem generalizes the insights of this example.

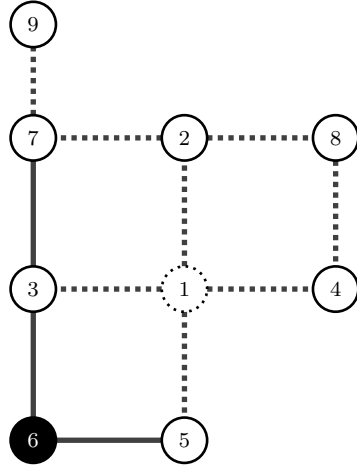
Theorem 3.4. Let $G = (V, E)$ be an undirected connected planar graph with $|E| \geq 2$. In an optimal solution to Epm , if edge (j, k) is assigned to center node i (i.e., $x_{i,(j,k)} = 1$), then there exists an optimal solution where all edges in $SP_{i,(j,k)}$ are also assigned to i .



(a) Road Network to be Partitioned into Two Compact and Contiguous Territories with E_pM .



(b) E_pM Solution before Imposing SPC (Dispersion = 35).



(c) E_pM Solution after Imposing SPC (Dispersion = 30).

— First Territory
 Second Territory

Figure 3: Two E_pM Solutions of the Road Network.

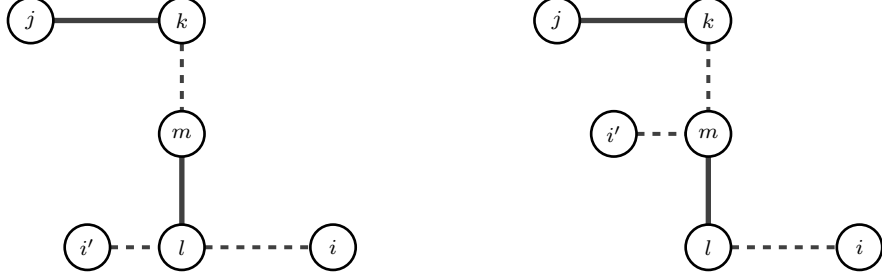
Proof. Proof. Assume by contradiction that (j, k) is assigned to i , but there is some edge along $SP_{i,(j,k)}$ not allocated to node i , and that this solution achieves a lower dispersion value than the solution where (j, k) is assigned to i and all edges along $SP_{i,(j,k)}$ are allocated to i . Without loss of generality, let $(l, m) \in SP_{i,(j,k)}$ be the edge that is instead allocated to center node $i' \neq i, j, k$, and let node l be the tail node of $SP_{i,(l,m)}$.

The only difference between the dispersion measure of the two solutions is related to the difference between $d_{i',(j,k)}$ and $d_{i,(j,k)}$, because the location and allocation variables associated with all other node-edge pairings have identical values. We consider two possible cases and accompany them with helpful illustrations (see Figure 4).

Case 1: l is the tail node $SP_{i',(l,m)}$ (Figure 4a).

Since $d_{i',(l,m)} < d_{i,(l,m)}$ holds, we have that

$$d_{i,(j,k)} = d_{i,(l,m)} + d_{l,(j,k)} \quad (13a)$$



(a) Planar Graph Example for Case 1. (b) Planar Graph Example for Case 2.

Figure 4: Planar Graph Examples to Illustrate Two Cases for the Proof of Theorem 3.4

$$> d_{i',(l,m)} + d_{l,(j,k)} \quad (13b)$$

$$= d_{i',(j,k)}. \quad (13c)$$

Equation (13a) corresponds to the distance from node i to edge (j, k) , whereas (13b) corresponds to the distance from i' to (j, k) . Since $d_{i',(l,m)} < d_{i,(l,m)}$ holds, then $d_{i,(j,k)} > d_{i',(j,k)}$. This contradicts the assumption that center node i is the optimal center for edge (j, k) .

Case 2: m is the tail node of $SP_{i',(l,m)}$ (Figure 4b).

In this case, since $d_{i',(l,m)} < d_{i,(l,m)}$ and $l_{(l,m)} > 0$ hold, we have that

$$d_{i,(j,k)} = d_{i,(l,m)} + d_{l,(j,k)} \quad (14a)$$

$$= d_{i,(l,m)} + l_{(l,m)} + d_{m,(j,k)} \quad (14b)$$

$$> d_{i',(l,m)} + d_{m,(j,k)} \quad (14c)$$

$$= d_{i',(j,k)}. \quad (14d)$$

Equation (14a) corresponds to the distance from node i to edge (j, k) , whereas (14c) corresponds to the distance from i' to (j, k) . Since $d_{i',(l,m)} < d_{i,(l,m)}$ and $l_{(l,m)} > 0$ hold, then $d_{i,(j,k)} > d_{i',(j,k)}$. This contradicts the assumption that node i is an optimal center for edge (j, k) . \square

3.4 Computational Results for ECpM

Next, we describe computational tests to evaluate the impact of the SPC constraints on solving the ECpM and EpM problems. The optimization models were coded in Python 3.8 and solved using CPLEX 22.1.1. The computational tests were performed on 28 cores of a high performance computing node equipped with 128 GB of memory and an Intel Xeon E5 v4 Processor running at 2.40 GHz. For each instance and solution method, a time limit of 12 hours was imposed.

The problem instances are derived from 14 real-world road networks located in Denmark, which were originally developed for the capacitated arc routing problem (Küilerich and Wøhlk, 2018); they are deployed therein for waste collection and are herein repurposed for more general location-allocation problems to test the featured EpM models. The selected road networks encompass small-, medium, and large-scale problem instances, with the cardinality of V ranging from 198 to 2,773 nodes and the cardinality of E ranging from 265 to 3,472 edges. From these fourteen road networks, a total of 84 instances are created by fixing one of six numbers of districts, namely, $p \in \{2, 10, 30, 40, 50, 100\}$.

Table 1 shows key instance information (the respective column labels are provided in the adjoining parentheses): road network number (RN), road network name (RN Name), number of nodes ($|V|$), number of edges ($|E|$), and the associated number of binary variables (NBV) in the $ECpM$ model. The table also shows computational performance metrics: average computational time of solving $ECpM$ with cut set contiguity constraints via the branch-and-cut scheme ($ECpM$ -CSC), average computational time of solving $ECpM$ with the shortest-path contiguity constraints ($ECpM$ -SPC), and the improvement factor of $ECpM$ -SPC relative to $ECpM$ -CSC (CSC/SPC). More specifically, the latter performance metric is calculated as

$$\text{CSC/SPC} = \frac{\text{Average time to solve the instances with } ECpM\text{-CSC}}{\text{Average time to solve the instances with } ECpM\text{-SPC}}. \quad (15)$$

Hence, when $\text{CSC/SPC} > 1$, this indicates that $ECpM$ -SPC was solved faster than $ECpM$ -CSC.

Table 1: On Average, $ECpM$ -SPC Outperformed $ECpM$ -CSC across All Tested Road Networks.

RN No.	RN Name	$ V $	$ E $	NBV	$ECpM$ -CSC	$ECpM$ -SPC	CSC/SPC
1	F15_g	198	265	52,668	142	8	17.31
2	N17_g	268	305	82,008	141	15	9.58
3	K13_g	393	421	165,846	284	34	8.24
4	F6_p	502	741	372,484	1,414	94	14.98
5	O12_g	761	852	649,133	1,766	202	8.74
6	S9_p	782	1,030	806,242	1,958	305	6.43
7	N16_g	915	1,030	943,365	2,086	404	5.16
8	S9_g	1,094	1,447	1,584,112	4,151	962	4.31
9	S11_g	1,564	1,805	2,824,584	7,005	2,985	2.35
10	K9_p	1,735	2,226	3,863,845	12,445	4,376	2.84
11	N11_g	2,134	2,411	5,147,208	18,987	6,778	2.80
12	N15_g	2,171	2,459	5,340,660	21,357	7,317	2.92
13	N9_g	2,018	2,655	5,359,808	18,622	7,131	2.61
14	K9_g	2,773	3,472	9,630,629	*	16,880	*

* The computer ran out of memory for all instances associated with this road network.

Analysis of the computational impact of the SPC constraints on the edge-based contiguous p -median problem. The first salient observation is that, for all six instances associated with the largest road network tested, namely RN14, $ECpM$ -CSC ran out of memory. Conversely, $ECpM$ -SPC was able to solve each instance associated with this road network in 4.69 hours, on average. Restricted to instances associated with RN1 to RN13, the average and median improvement factors of $ECpM$ -SPC relative to $ECpM$ -CSC (i.e., CSC/SPC) were 6.79 and 5.16, respectively. These improvement factors decrease as the network size increases, ranging from 17.31 for RN1 to 2.61 for RN13. Over these thirteen road networks, $ECpM$ -SPC outperformed $ECpM$ -CSC in 83.3% (65 of 78) of all associated instances. The remaining instances where $ECpM$ -SPC did not achieve the fastest computational times each involved small settings of the district parameter, namely, $p = 2$,

10. In other words, $ECpM$ -SPC tends to outperform $ECpM$ -CSC as the number of possible location and allocation decisions increases in the network.

The computational benefits of the SPC constraints are more pronounced on instances with more districts. As a representative example, for the instance of RN12 with $p = 100$, solving $ECpM$ -CSC took 10.88 hours, but solving $ECpM$ -SPC took only 54.7 *minutes*, yielding an 11.95x improvement. To further delve into the computational benefits of the SPC constraints as p increases, we focus on four road networks of increasing sizes, namely, RN1, RN5, RN9, and RN12. Figure 5 shows that CSC/SPC increases as p increases; here, note that the improvement factors are calculated using the solution times attained by $ECpM$ -CSC and $ECpM$ -SPC for each individual instance (as opposed to the average times over multiple instances). One plausible explanation for the observed increase is that, as the number of districts increases, branch-and-cut incurs more computational overhead—i.e., a higher number of variable subsets must be checked for contiguity and separating constraints added, whenever an integer-feasible solution is found. Conversely, $ECpM$ -SPC has a fixed polynomial number of contiguity constraints, more specifically $\mathcal{O}(|V||E|)$ of them, and these are always added in full up front.

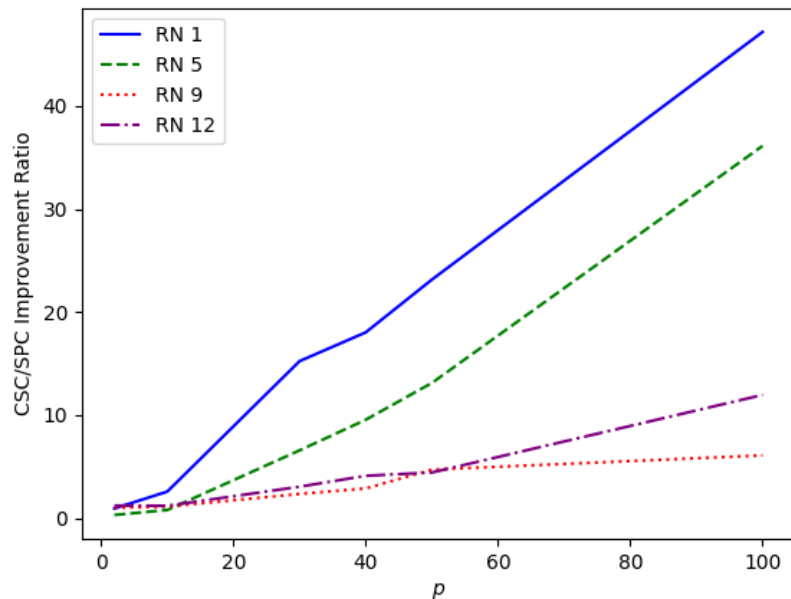


Figure 5: The improvement of $ECpM$ -SPC relative to $ECpM$ -CSC tends to grow with p

Analysis of the computational impact of the SPC constraints on the edge-based p -median problem. Next, we illustrate usefulness of the SPC constraints beyond enforcing contiguity. Specifically, we measure their effect on solving the E_pM problem—recall that the SPC constraints represent supervalid inequalities for this simpler problem, as is explained in Section 3.3. To that end, we compare the solution times of the E_pM and $ECpM$ -SPC models over the fourteen road networks; the results are summarized in Table 2.

For instances associated with RN1 to RN8, E_pM outperformed $ECpM$ -SPC, which suggests that the larger model size caused by the addition of the SPC constraints had a bigger

effect than the resulting reduction in the feasibility space. On the other hand, for all but one instance associated with RN9 to RN13 (each with $|E| \geq 1,805$), ECpM-SPC outperformed EpM by 2.45x, on average. Similar to ECpM-CSC in the prior experiment, EpM ran out of memory for all instances associated with RN 14, but ECpM-SPC solved all of these instances to optimality, which is surprising in that EpM is a relaxation of ECpM. This illustrates that the benefit of adding the SPC constraints as supervalid inequalities for EpM is best realized on larger and more complex problem instances.

Table 2: Adding the SPC Constraints as Supervalid inequalities of EPM Was Beneficial for Those Instances Associated with the Six Largest Road Networks.

RN No.	RN Name	$ V $	$ E $	NBV	EPM	ECpM-SPC	EPM/SPC
1	F15_g	198	265	52,668	5	8	0.64
2	N17_g	268	305	82,008	10	15	0.65
3	K13_g	393	421	165,846	19	34	0.55
4	F6_p	502	741	372,484	43	94	0.45
5	O12_g	761	852	649,133	56	202	0.27
6	S9_p	782	1,030	806,242	106	305	0.35
7	N16_g	915	1,030	943,365	126	404	0.31
8	S9_g	1,094	1,447	1,584,112	346	962	0.36
9	S11_g	1,564	1,805	2,824,584	5,497	2,985	1.84
10	K9_p	1,735	2,226	3,863,845	9,128	4,376	2.09
11	N11_g	2,134	2,411	5,147,208	14,913	6,778	2.20
12	N15_g	2,171	2,459	5,340,660	15,622	7,317	2.14
13	N9_g	2,018	2,655	5,359,808	13,696	7,131	1.92
14	K9_g	2,773	3,472	9,630,629	*	16,880	*

* The computer ran out of memory for all instances associated with this road network.

4 Connections to Edge-based Districting Models

ECpM can be transformed into a more traditional logistics districting model by imposing a work balance criterion, which is intended to promote fairness among different service agents (e.g., drivers) (Yanik and Bozkaya, 2020). This criterion can be imposed by requiring that the same quantity of demand must be served within each territory, within a certain tolerance; this translates into setting an upper and a lower bound on the cumulative demand served within a district. To introduce these balance constraints, let $b_{(j,k)}$ be the demand along edge (j,k) , \bar{b} be a fixed quantity representing an equal apportionment of demand per district (i.e., obtained by dividing total demand by p), and τ be the allowed percentage deviation from satisfying \bar{b} of demand in each district exactly. The set of work balance constraints is written as,

$$\sum_{(j,k) \in E} b_{(j,k)} x_{i,(j,k)} \leq \bar{b}(1 + \tau) w_i \quad \forall i \in V \quad (16)$$

$$\sum_{(j,k) \in E} b_{(j,k)} x_{i,(j,k)} \geq \bar{b}(1 - \tau) w_i \quad \forall i \in V. \quad (17)$$

Constraints (16) and (17) give upper and lower bounds on the total demand served within a territory. We remark that Constraints (4) become redundant when Constraints (16) are added, as the latter also fulfill the requirement of allowing edge (j, k) to be allocated to node i only if i is selected to be a center node. Thus, in the subsequent computational tests in Section 4.2, Constraints (4) are dropped from the featured districting formulations.

4.1 Differences between the Edge-based Districting Models Obtained with General Contiguity and Shortest-Path Contiguity

As the above paragraphs explain, adding Constraints (16) and (17) to $ECpM$ transforms it into an edge-based districting (EBD) problem. When they are added to $ECpM$ -CSC ((2) - (6), (7) - (8)), the model from Kassem and Escobedo (2023b) is obtained; the resulting model is denoted hereafter as EBD-CSC. However, a distinctive districting model is obtained when the balance constraints are added to $ECpM$ -SPC ((2) - (6), (11) - (12)); the resulting model is hereafter denoted as EBD-SPC. This stems from their interaction with shortest-path contiguity constraints, which enforce a stronger form of contiguity compared to the cut-set based constraints. Figure 6 provides a small example to illustrate the differences.

Example 4.1. Consider the connected planar graph depicted in Figure 6a, where $|V| = 11$, $|E| = 14$, $\tau = 20\%$, and $p = 2$; therein, it is assumed that $b_{(j,k)} = l_{(j,k)} \forall (j,k) \in E$, which gives $\bar{b} = 35.5$. Figures 6b and 6c depict the EBD-CSC and EBD-SPC solutions, respectively. In each case, dotted edges correspond to the first district, and solid edges correspond to the second district; the centers of the two districts follow the same formatting scheme as the edges. In both solutions, nodes 5 and 7 are the centers. The sole difference is the allocation of edge $(4, 9)$, which is allocated to the first district by EBD-CSC and to the second district by EBD-SPC. Note that allocating both $(1, 4)$ and $(4, 9)$ to 5 would violate the maximum demand allowed for the first district (42.6) and the minimum demand allowed for the second district (28.4). In effect, the combination of the SPC and balance constraints forces the allocation of $(4, 9)$ to the second district with center node 7 (since $(1, 4) \in SP_{7,(4,9)}$ and is allocated to the second district). Stated otherwise, these constraints disallow allocating $(4, 9)$ to the first district with center node 5, as SPC would force $(1, 4)$, which is part of $SP_{5,(4,9)}$, to be also allocated to the first district.

Although the dispersion values achieved by the optimal solutions of both models usually match, there are times when the EBD-CSC dispersion value can be lower than that of EBD-SPC. In fact, it is guaranteed that it will never be higher, since the latter imposes a stricter type of contiguity. Example 4.1 shows an instance with such a disparity, specifically, the dispersion value of EBD-CSC (55) is lower than that of EBD-SPC (60). The difference is due to the relationship between $d_{5,(4,9)}$ and $d_{7,(4,9)}$, namely, $10 = d_{5,(4,9)} < d_{7,(4,9)} = 15$. However, the lower dispersion attained by EBD-CSC comes with a practical drawback, namely that in some of the service areas, extra distance may be unnecessarily traversed by an agent in the subsequent routing operations, when one or more edges along the shortest path from a center node to one of its allocated edges (according to the original territory) is allocated to a different district. In Example 4.1, edge $(1, 4)$, which is part of $SP_{5,(4,9)}$, is not allocated to the first district. This means that to reach $(4, 9)$ from the center node 5, an agent assigned to serve the first district would have either to take a longer route

than the actual shortest path to stay within the district or temporarily leave the district, which could increase the deadhead distance (i.e., travel along roads where no service is performed (García-Ayala et al., 2016)). The differences between the two models also have computational implications, as the ensuing subsection demonstrates.

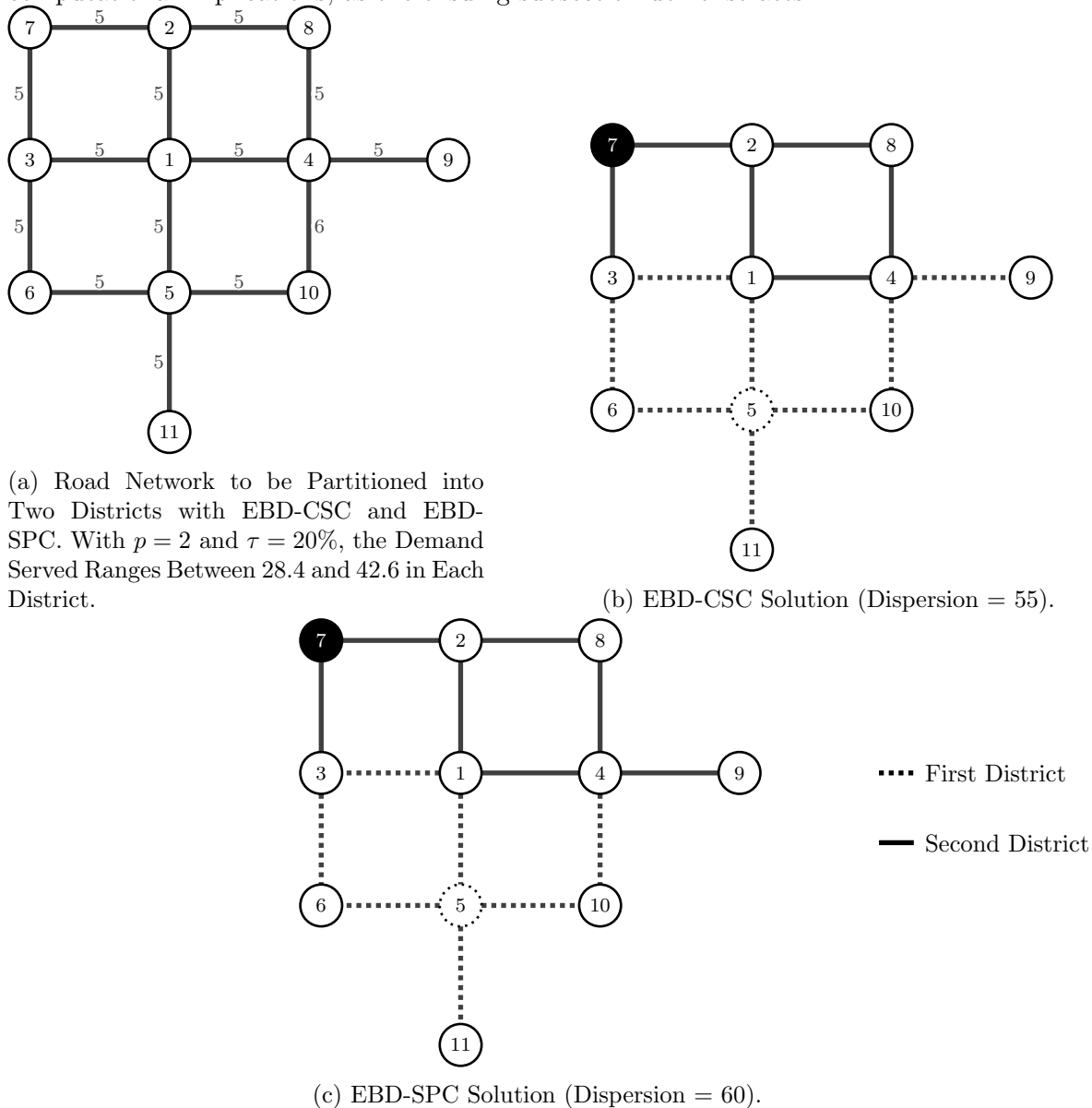


Figure 6: Road Network and Districting Solutions Obtained with EBD-CSC and EBD-SPC.

4.2 Computational Results for EBD

The same computational resources and specifications from Section 3.4 carry over to this section. The focus of the featured computational experiments is on EBD, which is a related but more challenging problem than $ECpM$. Owing to this increased difficulty, the computational tests feature only two medium-sized road networks, namely, RN 4 and RN 5. From

these two road networks, a total of 54 instances are created by fixing one of nine numbers of districts, namely, $p \in \{2, 4, 6, 8, 10, 20, 30, 40, 50\}$ and one of three work balance tolerance settings, namely, $\tau \in \{1\%, 10\%, 50\%\}$. In all of the generated instances, the demand along each edge (i.e., $b_{(j,k)}$) is set to its demand, as specified in Kiilerich and Wøhlk (2018).

The initial intent of the computational tests was to evaluate the different computational capabilities of the two districting models, namely, EBD-CSC and EBD-SPC. However, EBD-CSC could not find a feasible solution within the 12-hour time limit for any of the tested instances, meaning it did not reach a basic level of performance required for comparison. On the other hand, EBD-SPC was able to solve 59.3% (32 of 54) of the instances to optimality. Additionally, EBD-SPC reached the time limit for 16.7% (9 of 54) of the instances with an average relative percentage gap of 8.23%, it could not find a feasible solution within the 12-hour time limit for 3.70% (2 of 54) of these instances, and it returned infeasibility for 20.4% (11 of 54) of the instances. For these reasons, this section delves only into the computational results of EBD-SPC.

Analysis of the impact of the balance parameter τ . Different computational metrics were recorded to assess the effect of τ . They are recorded in Tables 3 and 4. Table 3 displays the road network number (RN No.), the tolerance setting (τ), the average computational time of solving EBD with the shortest-path contiguity constraints (EBD-SPC), the number of instances solved to optimality ($|\mathcal{I}_{\mathcal{T} \leq \infty} \neq \emptyset|$), the number of instances for which the time limit was reached with a feasible solution ($|\mathcal{I}_{\mathcal{T} > \infty} \neq \emptyset|$), the number of infeasible instances identified within the time limit ($|\mathcal{I}_{\mathcal{T} \leq \infty} = \emptyset|$), and the number of instances for which the time limit was reached without finding a feasible solution and without identifying the problem as infeasible ($|\mathcal{I}_{\mathcal{T} > \infty} = \emptyset|$). As a point of clarification, the average computational times of EBD-SPC are calculated based on instances that reached a feasible or optimal solution within the time limit (i.e., those instances belonging to $\mathcal{I}_{\mathcal{T} \leq \infty} \neq \emptyset$ or $\mathcal{I}_{\mathcal{T} > \infty} \neq \emptyset$). The road network number (RN No.), tolerance setting (τ), and average and max relative optimality gaps (Gap %) are reported in Table 4.

Tables 3 and 4 show that average computational times, average relative percentage gap, and maximum relative percentage gap tend to decrease as τ increases from 1% to 50%. This suggests that it is more difficult to balance the demand between the different districts at low tolerance settings. Indeed, 72.7% (8 of 11) of the infeasible instances obtained in this analysis arose from the setting $\tau = 1\%$. At this low value of τ , the demand must be balanced nearly equally among the districts, which is impossible for some instances.

Analysis of the impact of the number of districts. To test the sensitivity of p on computational times (in seconds) of solving EBD-SPC, for RNs 4 and 5, the average computational times were calculated over the three tolerance settings, $\tau \in \{1\%, 50\%, 100\%\}$, for each p setting, namely, $p \in \{2, 4, 6, 8, 10, 20, 30, 40, 50\}$. The results are plotted in Figure 7. For both road networks, there seems to be a general trend of decreasing average computational times as p increases. However, there are notable exceptions. For RN 5, there are slight increases at $p = 6$, $p = 10$, and $p = 40$; for RN 4, there are more sizable increases at $p = 6$, $p = 8$ and $p = 50$. One possible reason is that certain values of p can induce balance targets that are more difficult to achieve based on the demand values of a specific road network. Moreover, some of the variability could be explained by the fact that each plotted data point corresponds to the average of only three instances (i.e., one for each of the tested tolerance settings).

Table 3: Average EBD-SPC Solution Times (in Hours) Decrease as the Work Balance Tolerance (τ) Increases

RN No.	τ	EBD-SPC*	$ \mathcal{I}_{\mathcal{T} \leq \infty \epsilon} \neq \emptyset $	$ \mathcal{I}_{\mathcal{T} > \infty \epsilon} \neq \emptyset $	$ \mathcal{I}_{\mathcal{T} \leq \infty \epsilon} = \emptyset $	$ \mathcal{I}_{\mathcal{T} > \infty \epsilon} = \emptyset $
4	1%	10.20	2	3	4	0
4	10%	8.49	4	5	0	0
4	50%	0.84	9	0	0	0
5	1%	5.43	3	1	4	1
5	10%	2.08	5	0	3	1
5	50%	1.85	9	0	0	0

* Reported statistic includes instances where a feasible or optimal solution was reached within the 12 hour time limit (i.e., columns $|\mathcal{I}_{\mathcal{T} \leq \infty \epsilon} \neq \emptyset|$ and $|\mathcal{I}_{\mathcal{T} > \infty \epsilon} \neq \emptyset|$).

Table 4: Statistics on Relative Percentage Gap for Different Settings of τ in EBD-SPC

RN No.	τ	Gap %	
		Average	Max
4	1%	8.86%	12.31%
4	10%	4.94%	7.42%
4	50%	$\leq 0.01\%$	$\leq 0.01\%$
5	1%	22.85%	22.85%
5	10%	$\leq 0.01\%$	$\leq 0.01\%$
4	50%	$\leq 0.01\%$	$\leq 0.01\%$

5 Conclusion

This paper introduces and derives two binary programming formulations for the ECpM optimization problem. The first model requires an exponential number of cut set-based constraints and is paired with a branch-and-cut algorithm that generates only a small number of these constraints. The second utilizes a compact set of shortest-path contiguity (SPC) constraints derived herein; the model is solved with off-the-shelf methods. The results show that the SPC constraints can expedite the solution of large-scale instances of this problem and the simpler edge-based p -median problems (i.e., ECpM without contiguity), with speedups of up to 17x for the instances tested in this work. Finally, the paper explores structural insights and connections between ECpM and the edge-based districting (EBD) problem, which enforces an additional planning criterion (work balance). Balance constraints were added to the two ECpM models to obtain two EBD models. The shortest path-based formulation exhibited a superior computational performance, enabling the solution of larger instances, than the cut set-based formulation. As an additional practical benefit, the shortest path-based formulation eliminates a potential drawback in subsequent routing operations by preventing the buildup of deadhead distance and avoiding the need to travel beyond the district boundaries to reach specific customers. Future work will explore additional network insights that can further reduce the computational effort of solving ECpM and EBD.

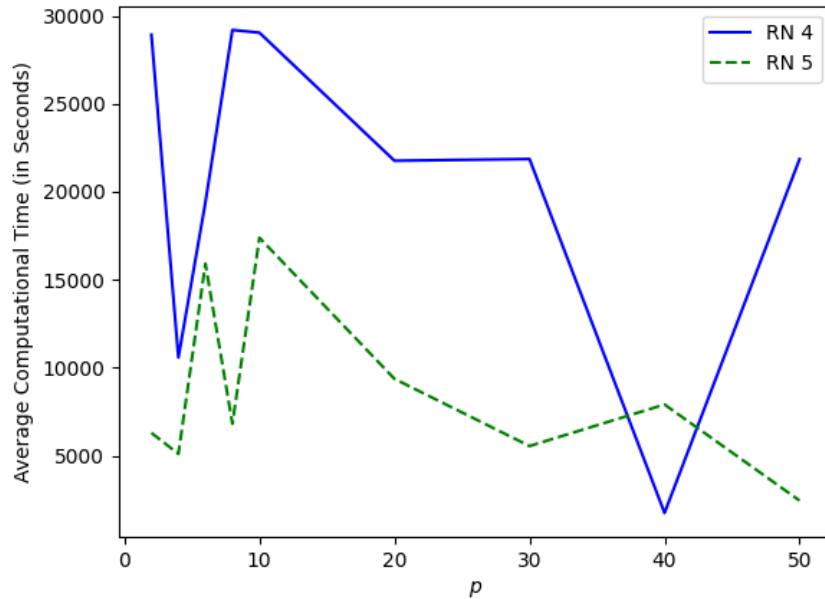


Figure 7: Effect of Increasing p on Average Computational Times of RN 4 and RN 5, with Different Tolerance Settings.

References

- Ahuja, R. K., T. L. Magnanti, and J. B. Orlin (1988). *Network Flows*. Prentice Hall, New Jersey.
- Alumur, S. A., J. F. Campbell, I. Contreras, B. Y. Kara, V. Marianov, and M. E. O’Kelly (2021). Perspectives on modeling hub location problems. *European Journal of Operational Research* 291(1), 1–17.
- Álvarez-Miranda, E. and J. Pereira (2021). A districting application with a quality of service objective. *Mathematics* 10(1), 13.
- Baycik, N. O. and K. M. Sullivan (2019). Robust location of hidden interdictions on a shortest path network. *IIE Transactions* 51(12), 1332–1347.
- Bergmann, F. M., S. M. Wagner, and M. Winkenbach (2020). Integrating first-mile pickup and last-mile delivery on shared vehicle routes for efficient urban e-commerce distribution. *Transportation Research Part B: Methodological* 131, 26–62.
- Bosona, T. (2020). Urban freight last mile logistics—challenges and opportunities to improve sustainability: a literature review. *Sustainability* 12(21), 8769.
- Branco, I. and J. Coelho (1990). The hamiltonian p -median problem. *European Journal of Operational Research* 47(1), 86–95.

- Brimberg, J., P. Hansen, N. Mladenovic, and S. Salhi (2008). A survey of solution methods for the continuous location-allocation problem. *International Journal of Operations Research* 5(1), 1–12.
- Butsch, A., J. Kalcsics, and G. Laporte (2014). Districting for arc routing. *INFORMS Journal on Computing* 26(4), 809–824.
- Cánovas, L., S. García, M. Labbé, and A. Marín (2007). A strengthened formulation for the simple plant location problem with order. *Operations Research Letters* 35(2), 141–150.
- Carvajal, R., M. Constantino, M. Goycoolea, J. P. Vielma, and A. Weintraub (2013). Imposing connectivity constraints in forest planning models. *Operations Research* 61(4), 824–836.
- Ceselli, A. and G. Righini (2005). A branch-and-price algorithm for the capacitated p-median problem. *Networks: An International Journal* 45(3), 125–142.
- Cova, T. J. and R. L. Church (2000). Contiguity constraints for single-region site search problems. *Geographical Analysis* 32(4), 306–329.
- Dablanc, L. and D. Rakotonarivo (2010). The impacts of logistics sprawl: How does the location of parcel transport terminals affect the energy efficiency of goods’ movements in paris and what can we do about it? *Procedia-Social and Behavioral Sciences* 2(3), 6087–6096.
- Dantzig, G., R. Fulkerson, and S. Johnson (1954). Solution of a large-scale traveling-salesman problem. *Journal of the operations research society of America* 2(4), 393–410.
- Daskin, M. S. (2013). *Median Problems*, Chapter 6, pp. 235–293. John Wiley Sons, Ltd.
- Drexl, A. and K. Haase (1999). Fast approximation methods for sales force deployment. *Management Science* 45(10), 1307–1323.
- Drezner, Z. and H. W. Hamacher (2004). *Facility location: applications and theory*. Springer Science & Business Media.
- Farughi, H., M. Tavana, S. Mostafayi, and F. J. Santos Arteaga (2020). A novel optimization model for designing compact, balanced, and contiguous healthcare districts. *Journal of the Operational Research Society* 71(11), 1740–1759.
- García-Ayala, G., J. L. González-Velarde, R. Z. Ríos-Mercado, and E. Fernández (2016). A novel model for arc territory design: Promoting eulerian districts. *International Transactions in Operational Research* 23(3), 433–458.
- Garfinkel, R. S. and G. L. Nemhauser (1970). Optimal political districting by implicit enumeration techniques. *Management Science* 16(8), B–495.
- Gollowitzer, S., D. L. Pereira, and A. Wojciechowski (2011). New models for and numerical tests of the hamiltonian p-median problem. In *International conference on network optimization*, pp. 385–394. Springer.

- Güden, H. and H. Süral (2019). The dynamic p-median problem with mobile facilities. *Computers & Industrial Engineering* 135, 615–627.
- Hakimi, S. (1979). An algorithmic approach to network location problems. ii: the p-median. *SIAM J. Appl. Math* 37, 539–560.
- Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations research* 12(3), 450–459.
- Hess, S. W., J. Weaver, H. Siegfeldt, J. Whelan, and P. Zitlau (1965). Nonpartisan political redistricting by computer. *Operations Research* 13(6), 998–1006.
- Hooker, J. N. (1994). Logic-based methods for optimization. In A. Borning (Ed.), *Principles and Practice of Constraint Programming*, Volume 874 of *Lecture Notes in Computer Science*, pp. 336–349. Springer, Berlin, Heidelberg.
- Israeli, E. and R. K. Wood (2002). Shortest-path network interdiction. *Networks: An International Journal* 40(2), 97–111.
- Jafari, N. and J. Hearne (2013). A new method to solve the fully connected reserve network design problem. *European Journal of Operational Research* 231(1), 202–209.
- Kalcsics, J. and R. Z. Ríos-Mercado (2019). Districting problems. *Location science*, 705–743.
- Kalcsics, J. and R. Z. Ríos-Mercado (2019). Districting problems. In G. Laporte, S. Nickel, and F. Saldanha da Gama (Eds.), *Location Science*, pp. 705–743. Springer, Cham, Switzerland.
- Kassem, Z. and A. R. Escobedo (2023a). The edge-based contiguous p-median problem: Formulations and path-based insights. In *IISE Annual Conference and Expo 2023*, pp. 96–101. Institute of Industrial and Systems Engineers, IISE.
- Kassem, Z. and A. R. Escobedo (2023b). Models and network insights for edge-based districting with simultaneous location-allocation decisions. *IISE Transactions* 55(8), 768–780.
- Külerich, L. and S. Wøhlk (2018). New large-scale data instances for carp and new variations of carp. *INFOR: Information Systems and Operational Research* 56(1), 1–32.
- Lolonis, P. and M. P. Armstrong (1993). Location-allocation models as decision aids in delineating administrative regions. *Computers, environment and urban systems* 17(2), 153–174.
- Marianov, V., D. Serra, and Z. Drezner (2002). Location problems in the public sector. *Facility location: Applications and theory* 1, 119–150.
- Martin, A. (2001). *General mixed integer programming: Computational issues for branch-and-cut algorithms*. Springer.

- Mayer, C., R. Mayer, M. A. Tariq, H. Geppert, L. Laich, L. Rieger, and K. Rothermel (2018). Adwise: Adaptive window-based streaming edge partitioning for high-speed graph processing. In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pp. 685–695. IEEE.
- Mayer, R. and H.-A. Jacobsen (2021). Hybrid edge partitioner: Partitioning large power-law graphs under memory constraints. In *Proceedings of the 2021 International Conference on Management of Data*, pp. 1289–1302.
- Mehrotra, A., E. L. Johnson, and G. L. Nemhauser (1998). An optimization based heuristic for political districting. *Management Science* *44*(8), 1100–1114.
- Melo, M. T., S. Nickel, and F. Saldanha-Da-Gama (2009). Facility location and supply chain management—a review. *European journal of operational research* *196*(2), 401–412.
- Mendes, L. H. P., F. L. Usberti, and C. Cavellucci (2022). The capacitated and economic districting problem. *INFORMS Journal on Computing*.
- Önal, H., Y. Wang, S. T. Dissanayake, and J. D. Westervelt (2016). Optimal design of compact and functionally contiguous conservation management areas. *European Journal of Operational Research* *251*(3), 957–968.
- Osorio, M. A. and M. Laguna (2003). Logic cuts for multilevel generalized assignment problems. *European Journal of Operational Research* *151*(1), 238–246.
- Osorio Lama, M. A. and R. Mújica García (1999). Logic cuts generation in a branch and cut framework for location problems. *Investigación Operativa* *8*(1-3), 155–166.
- Owen, S. H. and M. S. Daskin (1998). Strategic facility location: A review. *European journal of operational research* *111*(3), 423–447.
- Ricca, F., A. Scozzari, and B. Simeone (2011). Political districting: from classical models to recent approaches. *4OR* *9*, 223–254.
- Ricca, F. and B. Simeone (2008). Local search algorithms for political districting. *European Journal of Operational Research* *189*(3), 1409–1426.
- Salazar-Aguilar, M. A., R. Z. Ríos-Mercado, and M. Cabrera-Ríos (2011). New models for commercial territory design. *Networks and Spatial Economics* *11*(3), 487–507.
- Sefair, J. A., J. C. Smith, M. A. Acevedo, and R. J. Fletcher Jr (2017). A defender-attacker model and algorithm for maximizing weighted expected hitting time with application to conservation planning. *Iise Transactions* *49*(12), 1112–1128.
- Shirabe, T. (2009). Districting modeling with exact contiguity constraints. *Environment and Planning B: Planning and Design* *36*(6), 1053–1066.
- Tansel, B. C., R. L. Francis, and T. J. Lowe (1983). State of the art-location on networks: a survey. part i: the p-center and p-median problems. *Management science* *29*(4), 482–497.
- Ulukan, Z. and E. Demircioğlu (2015). A survey of discrete facility location problems. *International Journal of Industrial and Manufacturing Engineering* *9*(7), 2487–2492.

- Validi, H., A. Buchanan, and E. Lykhouyd (2022). Imposing contiguity constraints in political districting models. *Operations Research* 70(2), 867–892.
- Wang, D., J. Zhao, A. D’Ariano, and Q. Peng (2021). Simultaneous node and link districting in transportation networks: Model, algorithms and railway application. *European Journal of Operational Research* 292(1), 73–94.
- Wei, N. and J. L. Walteros (2024). On supervalid inequalities for binary interdiction games. *Mathematical Programming*, 1–42.
- Williams, J. C. (2002). A zero-one programming model for contiguous land acquisition. *Geographical Analysis* 34(4), 330–349.
- Wygonik, E. and A. V. Goodchild (2018). Urban form and last-mile goods movement: Factors affecting vehicle miles travelled and emissions. *Transportation Research Part D: Transport and Environment* 61, 217–229.
- Yanik, S. and B. Bozkaya (2020). *A Review of Districting Problems in Health Care*, pp. 31–55. Cham: Springer International Publishing.
- Yen, W. C.-K., S.-C. Chang, S. C.-S. Yang, and S.-C. Hu (2010). The connected p-median problem on graphs with forbidden vertices. In *Proceedings of 27th Workshop on Combinatorial Mathematics and Computation Theory, Providence University, Taichung, Taiwan*, pp. 168–173. Citeseer.
- Yuan, Y., D. Cattaruzza, M. Ogier, C. Rousselot, and F. Semet (2021). Mixed integer programming formulations for the generalized traveling salesman problem with time windows. *4OR* 19, 571–592.
- Yuan, Y., D. Cattaruzza, M. Ogier, and F. Semet (2020). A note on the lifted miller–tucker–zemlin subtour elimination constraints for routing problems with time windows. *Operations Research Letters* 48(2), 167–169.
- Zhang, C., F. Wei, Q. Liu, Z. G. Tang, and Z. Li (2017). Graph edge partitioning via neighborhood heuristic. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 605–614.
- Zoltners, A. A. and P. Sinha (1983). Sales territory alignment: A review and model. *Management Science* 29(11), 1237–1256.