

## Edge covering with continuous location along the network

Kayhan Alamatsaz<sup>a</sup>, Ali Aghadavoudi Jolfaei<sup>b</sup> and Mehdi Iranpoor<sup>c\*</sup>

<sup>a</sup>M.Sc. student, Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran

<sup>b</sup>M.Sc. student, Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan, Iran

<sup>c</sup>Assistant professor, Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan, Iran

### CHRONICLE

#### Article history:

Received February 5 2020  
Received in Revised Format  
February 28 2020  
Accepted March 5 2020  
Available online  
March 5 2020

#### Keywords:

Edge covering  
Unrestricted facility location  
Mathematical formulation  
Benders decomposition  
Matheuristic

### ABSTRACT

The set covering problem is to find the minimum cardinality set of locations to site the facilities which cover all of the demand points in the network. In this classical problem, it is assumed that the potential facility locations and the demand points are limited to the set of vertices. Although this problem has some applications, there are some covering problems in which the facilities can be located along the edges and the demand exists on the edges, too. For instance, in the public service environment the demand (population) is distributed along the streets. In addition, in many applications (like bus stops), the facilities are not limited to be located at the vertices (intersections), rather they are allowed to be located along the edges (streets). For the first time, this paper develops a novel integer programming formulation for the set covering problem wherein the demand and facility locations lie continuously along the edges. In order to find good solutions in a reasonable time, a matheuristic algorithm is developed which iteratively adds dummy vertices along the edges and solves a simpler problem which does not allow non-nodal facility locations. Finally, a Benders decomposition reformulation of the problem is developed and the lower bounds generated by the Benders algorithm are used to evaluate the quality of the heuristic solutions. Numerical results show that the Benders lower bounds are tight and the matheuristic algorithm generates good quality solutions in short time.

© 2020 by the authors; licensee Growing Science, Canada

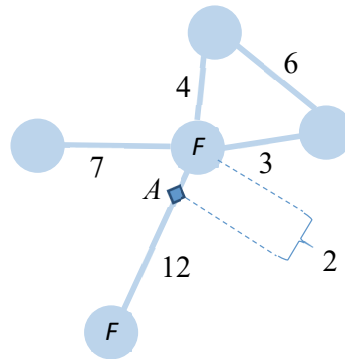
## 1. Introduction

Placing facilities in appropriate locations has a long history but it has been received a great deal of formal researches since 1950s (Heragu, 2008). The set covering problem is one of the most well-known facility-location problems with many applications, especially in the public sector (Farahani et al., 2012; Snyder, 2011). In this problem, a demand point would be covered if its distance from at least one sited facility does not exceed a given length. The objective of the set covering problem is the minimization of the number of used facilities to cover all of the demand points (Snyder, 2011). This problem was introduced by (Hakimi, 1965). He addressed the problem of finding the minimum number and the locations of policemen to cover the vertices of a highway network. Toregas et al. (1971) proposed the famous integer programming formulation for this problem.

\* Corresponding author Tel.: +983133911485; Fax: +983133915526  
E-mail: [iranpoor@cc.iut.ac.ir](mailto:iranpoor@cc.iut.ac.ir) (M. Iranpoor)

The abovementioned classical set covering problem assumes that the facilities could be located only at the vertices of the network and the demand occurred only at the vertices (Toregas et al., 1971). While these two assumptions have their own applications, there are other circumstances where the facilities have insignificant geographic footprint and hence can be located anywhere along the network (Wei et al., 2014); and at the same time, the demand for these facilities lies continuously along the edges. For example, we can refer to bus stops, interactive kiosks, ATMs, bike sharing stations, convenience shops, car charging stations, footbridges and wildlife road crossings, surveillance cameras, municipal waste collection bins, accident reporting acoustic sensors on road networks, and fire hydrants, some of which are mentioned in Berman et al. (2016) and Wei et al. (2014). As compared to the classical variation of the problem, there are two implications in dealing with this unrestricted variation of the problem (Church & Murray, 2018). Firstly, an infinite number of candidate facility locations are to be regarded. Secondly, there is not a finite set of demand points anymore; instead the continuous parts of the network (i.e. edges) must be considered for coverage.

Hakimi property (also known as nodal optimality) states that in the  $p$ -median problem, there is always at least one optimal solution that the facilities are sited only on the vertices (Laporte et al., 2015). It is known that this property does not carry over for the classical set covering problem (Snyder, 2011). Here, by using a simple counterexample, we show that this property does not hold for the case in which the demand is distributed along the edges. Hence, the relaxation of the restriction that the facilities must be located at the vertices can decrease the number of required facilities to cover the network. Fig. 1 represents a small network with coverage distance 10 in which the number beside each edge is its length. It can be verified that if the facilities are restricted to be located at vertices, the network is covered by at least two facilities (e.g. the ones denoted by  $F$ ). However, the relaxation of this restriction allows the coverage of the whole network by only one facility located at point  $A$ .



**Fig. 1.** Unrestricted location of facilities

In Fig. 1, the coverage of the network is equivalent to the coverage of its vertices. However, this is not a general rule. Consider Fig. 2 with coverage distance 10. Again, the length of each edge is written beside it and the located facilities are denoted by  $F$ . All vertices are covered by these facilities, but some part of the network (i.e. the dotted section) has been left uncovered. Hence, the coverage of the vertices is not necessarily equivalent to the coverage of the entire network.



**Fig. 2.** Total Coverage of the vertices with a partially uncovered edge

Similar to the edge covering problem, the *minimum vertex cover problem* is to find the minimum-size set of vertices which covers all of the edges (Cormen et al., 2009). However, the notion of coverage is different and is not based on the coverage distance, rather it is assumed that each edge is covered if a facility is located on one of its endpoints or both. Some researchers like Li et al. (2016) and Cai et al. (2019) addressed the weighted extension of this problem in which the aim is to minimize the total weights

of the selected vertices. They designed some efficient local search algorithms to solve this problem. Pandey and Punnen (2018) studied a generalization of the minimum vertex cover problem in which the objective function is composed of the costs of not covering, covering, and double covering of edges. They developed the integer programming formulations of the problem and analyzed its computational complexity.

The formal discussion on the edge covering problem regarding the coverage distance dates back to 1976 when ReVelle et al. studied the set covering problem with the edge demand. They assumed that an edge is called covered if at least one sited facility reaches all of its points, meaning that the coverage of one edge cannot be accomplished by the joint partial coverage from multiple facilities. Sadigh et al. (2010) extended the edge covering problem allowing each edge to be covered from two directions by partial coverage of two facilities. They solved the problem using a Tabu search algorithm equipped with size reduction rules which decrease the solution time. Similar to ReVelle et al. (1976), they assumed that the candidate sites to place the facilities are limited to the set of vertices. Conrow et al. (2018) studied the bike sharing siting problem in the city of Phoenix. They considered two objectives of maximization of the coverage of demand points and edge segments. They considered a finite set of potential locations for bike sharing stations.

Covering problems with facility location on the edges and discrete demand points have also been studied. Church and Meadows (1979) addressed the maximal and set covering problems wherein the demands exist on the vertices but the facilities are allowed to be sited anywhere along the network. For both problems, they proved that there is an optimal solution that the facilities are located at vertices plus some specific points along the edges. They also formulated the set covering and maximal covering of edges with the restriction that the facilities can be located only on the vertices and each edge is allowed to be covered by multiple facilities. Groß et al. (2009) tackled the problems of locating new facilities (e.g. bus stops and railway stations) along the edges of a transportation network in order to satisfy either the demand of some finite points in the plane or the nodes of the network. They considered two objectives, separately: minimizing the number of stations which cover all demand points and minimizing the total distance between demand points and their nearest facilities when the number of stations is given. Also, they proved that there are finite sets of optimal location points of polynomial order. Schobel et al. (2009) studied adding new bus stops (or train stations) to an existing bus route (or railway network) with the objective of minimization of total added stopping times subject to the constraint that all of the discrete set of demand points must be covered. They converted the problem to the classical set covering problem and found some cases that can be solved more efficiently. Jayalakshmi et al. (2017) addressed cooperative maximum covering problem wherein the coverage was not binary, rather its strength decreases linearly with increase in distance. Further, they assumed that each demand point is considered as covered if the sum of partial coverage from all facilities reaches a threshold. They assumed that the sites for locating facilities include continuous parts of edges as well as the vertices.

A few studies have considered the facility location problems when both demand and location occur continuously along the network. De Los Mozos & Mesa (2000) tackled the single facility location problem along the network wherein the demand lies both on the vertices and along the edges. By knowing the density function of the level of edge demand, they developed a decomposition algorithm to minimize the variance of distance between the located facility and all demand points. They showed that this approach finds the optimal location of the facility in a more efficient manner than the exhaustive approach. Berman et al. (2016) tackled the maximal covering problem that the demand is distributed along the edges and facilities can be located anywhere on the network. For the case when the demand is distributed constantly along edges and there is only one facility to be located on the network, they derived the discrete set of locations called finite dominating set (*FDS*) which always contains the optimal location. Based on the *FDS*, they designed a greedy adding and improvement heuristic to solve the multi-facility location problem. They also addressed the minimal covering problem of obnoxious facilities. Wei et al. (2014) addressed the set covering problem where the facilities can be located anywhere along the network and the demand exists all over the edges. They designed an algorithm which generates lower

and upper bounds iteratively in which the gap between these bounds approaches zero as the algorithm proceeds. Their algorithm was based on the assumption that the coverage distance is defined based on the Euclidian distance (i.e. each demand point is covered if at least one facility is located in its coverage circular disk).

The contributions of this paper are as follows. First, this paper tackles a novel extension of the set covering problem where the facilities can be located continuously along the network and the aim is to cover all of the edges (i.e. the whole network) with the minimum number of facilities. To the best of our knowledge, the only study which has considered this problem is Wei et al. (2014). However, Wei et al. (2014) assumed that the coverage is based on the Euclidian distance of the points on the plane. On the contrary, we assume that the coverage is defined based on the shortest path along the network. This seems more practical for public facilities in which the customers (facilities) should pass through the network to reach the facilities (customers). Second, a novel mathematical programming formulation of the problem is developed for the first time. Third, a matheuristic algorithm is designed which can find near optimal solutions in a short time. Fourth, a Benders algorithm is developed to generate lower bounds which verify the quality of the matheuristic solutions.

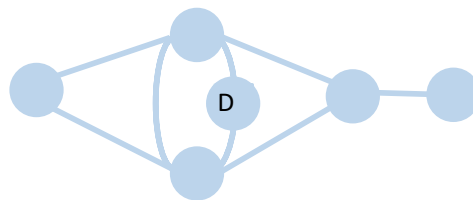
The remaining parts of the paper are organized as follows. Section 2 introduces the problem and develops its mathematical formulations. Section 3 is dedicated to the solution approaches. Subsection 3.1 provides some dominance rules. Subsection 3.2 introduces the node-adding matheuristic approach. Subsection 3.3 presents the Benders decomposition formulation of the problem which is used to generate the lower bound solutions. Section 4 uses the numerical results to verify the quality of the matheuristic algorithm and the tightness of the Benders' lower bounds. Finally, Section 5 concludes the paper and sheds light on some future extensions.

## 2. Problem formulation and notations

This paper tackles a novel edge covering problem wherein the facilities can be located anywhere on the network. The objective is to determine the minimum number of required facilities and their positions on the network. An edge does not need to be covered by a single facility, rather it can be covered cooperatively. The following assumptions are considered in the sequel.

- The graph of the problem is connected and undirected.
- At most one facility can be located along each edge.
- There is no parallel edges between any pair of vertices.
- All edges must be covered.
- No edge is longer than twice of the coverage distance.

It should be stated that the multi-graphs can be handled by the following simple trick. As shown in Fig. 3, adding a dummy vertex (like  $D$ ) on an arbitrary position of each parallel edge converts the multi-graph to an ordinary graph.



**Fig. 3.** Adding the dummy vertices to handle the parallel edges

The last assumption guarantees the feasibility of the problem. Although the violation of this assumption does not necessarily lead to infeasibility of the problem, the feasibility can be assured by adding some dummy vertices over the edges which are longer than two times of the coverage distance.

The sets and the parameters of the problem are as follows.

$V$ : set of vertices (also called nodes)

$E$ : set of edges

$n$ : the number of vertices

$m$ : the number of edges

$D_c$ : the coverage distance

$l_{ij}$ : the length of edge  $ij$

$M$ : a sufficiently large number

$d_{ij}$ : the length of the shortest path between vertex  $i$  and vertex  $j$

Further, the decision variables are denoted by the following notations:

$x_{ij}$ : binary variable which is 1 if a facility is located on edge  $ij$  ( $i < j$ ) and 0 otherwise

$y_{ij}$ : the distance of the located facility on edge  $ij$  from vertex  $i$

$c_{ij}$ : the maximum covered length of edge  $ij$  through vertex  $i$

$\delta_{iji'}$ : binary variable which is 1 when the coverage of edge  $i'j'$  from vertex  $i'$  is done by the located facility on edge  $ij$  through the shortest path between  $i$  and  $i'$  and 0 otherwise

$e_{ij}$ : The covered length of edge  $ij$  from the located facility on this edge toward vertex  $i$

In the notations,  $ij$  and  $ji$  refer to the same edge. However, the order that nodes of the edges appeared in the indices of the variables  $y_{ij}$ ,  $c_{ij}$ ,  $\delta_{iji'}$ ,  $e_{ij}$  refers to the direction in which the edge is considered. An example of the complementary coverage of an edge ( $ij$ ) with two facilities (denoted by diamond) is depicted in Figure 4.

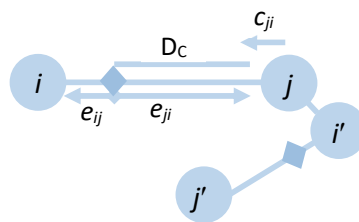


Figure 4. The exemplification of the complementary edge covering

The parameters  $d_{ij}$ s are determined before solving the problem by running a shortest path algorithm for every pair of vertices. To this end, we used Dijkstra's algorithm (Cormen et al., 2009) and implemented it in Matlab. The mathematical formulation of the problem is as follows:

$P$ :

$$\min \sum_{ij \in E | i < j} x_{ij} \quad (1)$$

$$y_{ij} + y_{ji} = l_{ij}; \quad ij \in E | i < j \quad (2)$$

$$\sum_{i'j' \in E | i' \neq ij} \delta_{ij i'j'} \leq 2(m-1)x_{ij} ; ij \in E | i < j \quad (3)$$

$$\sum_{i'j' \in E | i' \neq ij} \delta_{ij i'j'} \leq 2(m-1)x_{ji} ; ij \in E | i > j \quad (4)$$

$$\sum_{ij \in E | ij \neq i'j'} \delta_{ij i'j'} \leq 1 ; i'j' \in E \quad (5)$$

$$c_{i'j'} \leq D_C - y_{ij} - d_{i'j'} + M(1 - \delta_{ij i'j'}) ; i'j', ij \in E \& ij \neq i'j' \quad (6)$$

$$c_{i'j'} \leq l_{i'j'} \sum_{ij \in E | ij \neq i'j'} \delta_{ij i'j'} ; i'j' \in E \quad (7)$$

$$e_{ij} \leq D_C x_{ij} ; ij \in E | i < j \quad (8)$$

$$e_{ij} \leq D_C x_{ji} ; ij \in E | i > j \quad (9)$$

$$e_{ij} + c_{ij} \leq y_{ij} ; ij \in E \quad (10)$$

$$c_{ij} + c_{ji} + e_{ij} + e_{ji} \geq l_{ij} ; ij \in E | i < j \quad (11)$$

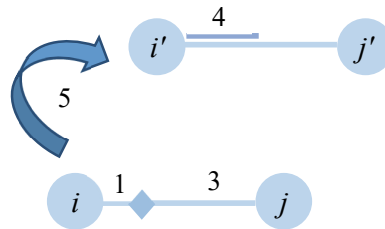
$$x_{ij} \in \{0,1\} ; ij \in E | i < j \quad (12)$$

$$y_{ij}, e_{ij}, c_{ij} \geq 0 ; ij \in E \quad (13)$$

$$\delta_{ij i'j'} \in \{0,1\} ; ij, i'j' \in E \& ij \neq i'j' \quad (14)$$

The objective function (1) minimizes the number of used facilities. Constraint (2) states that the sum of the distance of the location of any facility from the endpoints of its edge equals the length of that edge. We remind that for the edges without a located facility, the value of variable  $y$  has no meaning. Constraints (3) and (4) ensure that if at least one edge is covered by edge  $ij$ , a facility have to be located along it. Constraint (5) imposes that edge  $i'j'$  is covered through vertex  $i'$  by at most one located facility along the other edges.

Constraints (6) and (7) determine the length of edge  $ij$  which is covered by a located facility along other edges. To clarify these two constraints, consider the example depicted in Fig. 5 where a facility (denoted by a diamond) is located along edge  $ij$ . Suppose that the coverage distance ( $D_C$ ) is 10. Assume that the shortest path from node  $i$  to node  $i'$  is of length 5 (i.e.  $d_{ii'}=5$ ). As shown in this figure,  $y_{ij}=1$ . Constraint (6) does not enforce  $\delta_{ij i'j'}$  to be zero. If  $\delta_{ij i'j'}=1$ , then  $c_{i'j'}$  can be 4. Now, assume that  $D_C$  is 4. In this case,  $\delta_{ij i'j'}$  has to be zero. Otherwise, Constraint (6) is violated. If for all edges  $i'j'$  we have  $\delta_{ij i'j'}=0$ , then Constraint (7) imposes that  $c_{i'j'}=0$ .



**Fig. 5.** An explanatory example for the calculation of  $c_{i'j'}$

Constraints (8-9) ensure that the covered length of edge  $ij$  by its own facility is not greater than the coverage distance. Constraint (10) prevents erroneous calculation of  $c_{ijs}$  by precluding twice consideration of edge parts. It is worth noting that when an edge does not contain a located facility, this constraint is redundant since the values of  $y_{ij}$  can take arbitrary values. Constraint (11) enforces every edge to be totally covered. The remaining constraints determine the domain of the variables.

It is worth noting that formulation  $P$  can also handle the graphs with directed edges. Suppose that edge  $ij$  is directed from  $i$  to  $j$ . Then, it suffices to drop all variables and constraints which consider traversing edge  $ij$  from  $j$  to  $i$ .

### 3. Solution approach

In this section first, some dominance rules are presented which reduce the feasible region. Further, a matheuristic algorithm is developed which solves a restricted variation of the problem, iteratively. Moreover, to evaluate the quality of the matheuristic solutions, Benders decomposition formulation of the problem is used to generate the benchmark lower bounds.

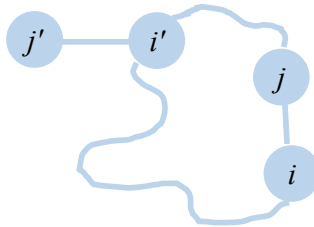
#### 3.1 Dominance rules

In this subsection, some dominance rules are explained with the aim of reducing the feasible region.

*Observation 1.* For any pair of nodes  $i$  and  $i'$ , if  $d_{ii'} > D_C$ , then  $\delta_{ij i'j} = 0$ . This is clear since any located facility along edge  $ij$  cannot cover any part of edge  $i'j'$  through the shortest path  $i \rightarrow i'$ .

*Observation 2.* For any pair of edges  $ij$  and  $i'j'$ , if  $d_{ii'} = d_{ji'} + l_{ij}$ , then  $\delta_{ij i'j} = 0$ . In these situations, regardless of the value of  $y_{ij}$ , the shortest path from the located facility along edge  $ij$  to node  $i'$  passes through node  $j$ . Hence, solutions with  $\delta_{ij i'j} = 1$  are dominated.

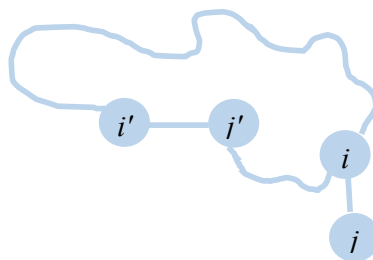
To illustrate this observation, consider the network shown in Figure 6 in which the straight lines represent the edges and the curves denote the shortest paths if the network excludes edge  $ij$ . Since  $d_{ii'} = d_{ji'} + l_{ij}$ , the curve  $ii'$  is longer than  $d_{ji'} + l_{ij}$ . It follows that the coverage of edge  $i'j'$  through node  $i'$  by a facility located along edge  $ij$  is always longer (or at least equal) if the coverage is considered through node  $j$ .



**Fig. 6.** Illustration of observation 2

*Observation 3.* For any pair of edges  $ij$  and  $i'j'$ , if  $d_{ii'} = d_{ij} + l_{i'j'}$ , then  $\delta_{ij i'j} = 0$ . In this case, the shortest path from node  $i$  to node  $i'$  traverses through node  $j'$ . Hence, solutions with  $\delta_{ij i'j} = 1$  are dominated.

To illustrate this observation, consider the network shown in Fig. 7 in which the straight lines represent the edges and the curves denote the shortest paths if the network excludes edge  $i'j'$ . Since  $d_{ii'} = d_{ij} + l_{i'j'}$ , the curve  $ii'$  is longer than  $d_{ij} + l_{i'j'}$ . It follows that if the coverage of edge  $i'j'$  by a facility that is located along edge  $ij$  through nodes  $i$  and  $i'$  is positive, this facility completely covers edge  $i'j'$  through nodes  $i$  and  $j'$ . Thus, the solutions with  $\delta_{ij i'j} = 1$  are dominated.



**Fig. 7.** An illustration of observation 3.

*Observation 4.* Suppose that  $ij$  is a pendant edge (that is,  $i$  or  $j$  has vertex degree 1) with  $l_{ij} \geq D_C$ . Without loss of generality, let  $i$  be the vertex with degree 1 and assume that  $i < j$ . Then, in any optimal solution  $x_{ij} = 1$ . Further, there is an optimal solution in which  $y_{ij} = D_C$ .

### 3.2 The Edge-division matheuristic approach (EDA)

The infinite number of candidate location points (along the edges) is one of the major sources of complexity of the problem  $P$ . The mathematical formulation of the restricted variation of the problem (we denote it by  $RV$  for short) in which the facilities are limited to be located on the vertices is as follows.

$RV$ :

$$\min \sum_{k \in V} x_k \quad (15)$$

$$\sum_{k \in V} c_{kij} \delta_{kij} + \sum_{k \in V} c_{kji} \delta_{kji} \geq l_{ij} ; ij \in E, i < j \quad (16)$$

$$\delta_{kij} \leq x_k ; k \in V, ij \in E \quad (17)$$

$$\sum_{k \in V} \delta_{kij} = 1 ; ij \in E \quad (18)$$

$$x_k \in \{0,1\} ; k \in V \quad (19)$$

$$\delta_{kij} \geq 0 ; k \in V, ij \in E \quad (20)$$

The objective is to minimize the number of located facilities. Parameters  $c_{kij}$  determine the coverable length of edge  $ij$  through vertex  $i$  if a facility is located at vertex  $k$ .  $c_{kij}$  is calculated beforehand and equals  $\min\{\max\{0, D_C - d_{ki}\}, l_{ij}\}$ . Regarding (18) and (20), the first summation of constraint (16) calculates the convex combination of scalars  $c_{kij}$  s. Thus, the maximum value of this expression for edge  $ij$  is resulted when  $\delta_{kij}$  equals 1 for  $k$  with maximum  $c_{kij}$  and zero for  $k' \neq k$ . This discussion holds for the second summation of constraint (16). Suppose these two expressions are in their maximum values. Now, if the sum of the covered lengths through both directions of edge  $ij$  is greater than or equal to the length of this edge, edge  $ij$  is covered completely.

The optimal solution of the problem  $P$  can be approximated by adding dummy vertices along the edges and solving problem  $RV$  which is easier. Adding more dummy vertices may improve the quality of the approximate solution but increases the time it takes to solve the problem. Assume that  $\Lambda$  is the number of dummy vertices which are supposed to be located along the edges. Figure 8 describes our proposed algorithm to distribute the dummy vertices evenly over the network.

```

Initialization:  $l'_{ij} \leftarrow l_{ij}, \gamma_{ij} \leftarrow 0; ij \in E$ 
For  $iter = 1$  to  $\Lambda$ 
    Find the edge in  $E$  with maximum  $l'_{ij}$  and call it as  $ij$ 
     $\gamma_{ij} \leftarrow \gamma_{ij} + 1$ 
     $l'_{ij} \leftarrow l'_{ij} \times \gamma_{ij} / (\gamma_{ij} + 1)$ 
For each edge  $ij \in E$  with  $\gamma_{ij} > 0$ ,
    Place  $\gamma_{ij}$  nodes along the edge so that the edge is divided into  $\gamma_{ij} + 1$  equidistant segments.

```

**Fig. 8.** Spreading the dummy vertices

The vertex set of the new graph includes the original and the dummy vertices. Moreover, the set of edges of this graph is composed of the undivided edges and the segments of the divided edges. The length of segments are  $l'_{ij}$ . Knowing  $l'_{ij}$ , the length of the shortest path between any pair of vertices which includes at least one dummy vertex can be easily calculated. The pseudocode of the edge-division matheuristic is presented in Fig. 9.



```

Initialization:  $RemT \leftarrow T, \lambda \leftarrow 0, G' \leftarrow$  Initial Graph
Solve problem  $RV$  by calling  $CPLEX$  with run time limit  $RemT$  for the graph with no dummy vertices
 $BestObj \leftarrow Obj$ 
 $BestSolution \leftarrow Solution$ 
 $RemT \leftarrow RemT - RunTime$ 
While  $RemT > 0$ 
    Augment  $G'$  with new graph by adding  $A=2^\lambda m$  dummy nodes distributed according to Fig. 8.
     $\lambda \leftarrow \lambda + 1$ 
    Solve problem  $RV$  by calling  $CPLEX$  for the new graph with run time limit of  $RemT$ 
     $RemT \leftarrow RemT - RunTime$ 
    If  $Obj < BestObj$ 
         $BestObj \leftarrow Obj$ 
         $BestSolution \leftarrow Solution$ 
Report  $BestObj$  and  $BestSol$ 

```

**Fig. 9.** The edge-division matheuristic

### 3.3 The Benders decomposition approach (BD)

In order to evaluate the performance of the heuristic solutions, we need their corresponding optimal or lower bound solutions. Problem  $P$  is a mixed integer programming problem with a huge number of constraints and binary variables which cannot be solved optimally by the general purpose solvers in a reasonable time except for small problems. Such solvers (like  $CPLEX$ ) also provide lower bound solutions, but these bounds are not necessarily tight. Benders technique is an optimization algorithm which decomposes the original problem into two smaller problems called master problem and subproblem (Taskin, 2010). This approach can solve problems with large scale formulations efficiently by alternately solving the master and subproblem (Oğuz et al., 2018). The master problem includes only a few (or even no) constraints. We use Benders decomposition algorithm to find tight lower bounds. In our case, the master problem is feasible and bounded, the subproblem is a feasibility problem, and the dual subproblem is also feasible. The optimal solution of  $P$  could be obtained by the solution of the master problem if it satisfies the constraints of the subproblem. Otherwise, the dual subproblem is unbounded and a constraint called feasibility cut needs to be appended to the master problem to exclude the corresponding extreme ray. The master problem ( $MP$ ) of problem  $P$  is as follows.

$MP$ :

$$\min \sum_{ij \in E | i < j} x_{ij} \quad (21)$$

$$\sum_{ij \in E | ij \neq i'j'} \delta_{ij'j'} \leq 1 ; i'j' \in E \quad (22)$$

$$\sum_{i'j' \in E | i'j' \neq ij} \delta_{ij'j'} \leq 2(m-1)x_{ij} ; ij \in E | i < j \quad (23)$$

$$\sum_{i'j' \in E | i'j' \neq ij} \delta_{ij'j'} \leq 2(m-1)x_{ji} ; ij \in E | i > j \quad (24)$$

$$x_{ij} \in \{0,1\} ; ij \in E | i < j \quad (25)$$

$$\delta_{ij'j'} \in \{0,1\} ; ij, i'j' \in E \& ij \neq i'j' \quad (26)$$

Here, we provide three initial cuts ( $IC$ ) with the aim of reducing the search space of the master problem.

Theorem 1. Expression (27) provides a lower bound for the number of required facilities.

$$F = \begin{cases} 1 & ; \max_{i,j \in E'} \{d_{ij}\} \leq 2D_C \\ \left\lceil \frac{\max_{i,j \in E'} \{d_{ij}\} - 2D_C - \varepsilon}{2D_C} \right\rceil + 2 & ; \max_{i,j \in E'} \{d_{ij}\} > 2D_C \end{cases} \quad (27)$$

where  $\varepsilon$  is an infinitesimal positive number and  $\lfloor x \rfloor$  is the integral part of  $x$ .

Proof.

The number of required facilities to cover the shortest path between the farthest nodes is a lower bound on the minimum number of required facilities to cover the entirety of the graph. First assume that the length of the shortest path between the farthest nodes is less than or equal to  $2D_C$ . This path can be covered by locating one facility at the middle of it. Now, consider the opposite case. Let the shortest path between the farthest nodes  $i$  and  $j$  be of length  $K+2D_C$ . The minimum number of required facilities to cover this path happens when two facilities are located at distance  $D_C$  from nodes  $i$  and  $j$ . If  $K$  is multiple of  $2D_C$ , the part of this path between these facilities can be covered by at least  $(K/(2D_C))-1$  facilities located along it. On the other hand, if  $K$  is not divisible by  $2D_C$ , the minimum number of required facilities to cover this path is attained by locating  $\lfloor K/2D_C \rfloor$  along this path.  $\square$

IC1. The following cut reduces the search region.

$$\sum_{ij \in E | i < j} x_{ij} \geq F \quad (28)$$

IC2. If edge  $i'j'$  does not accommodate a facility, its outer covering facilities need to be located along the edges that allow the complete coverage.

$$\sum_{ij \in E | ij \neq i'j' \& D_C - d_{ii'} > 0} (D_C - d_{ii'}) \delta_{ij i'j'} + \sum_{ij \in E | ij \neq i'j' \& D_C - d_{jj'} > 0} (D_C - d_{jj'}) \delta_{ij i'j'} + l_{ij} x_{ij} \geq l_{i'j'} ; i'j' \in E | i' < j' \quad (29)$$

According to this constraint, as shown in Figure 10, if two outer facilities (on edges  $ij$  and  $i''j''$ ) cannot completely cover edge  $i'j'$  when their facilities are located on the nearest place to this edge (that is, node  $i$  and node  $i''$ ), they cannot cover this edge regardless of the position of the located facilities.

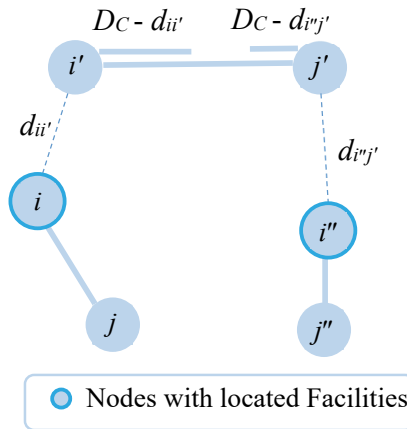
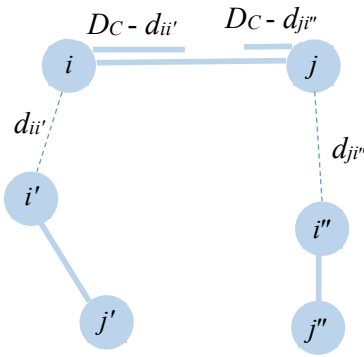


Fig. 10. Incomplete coverage

IC3.

$$\delta_{ij i'j'} + \delta_{jj'' i'j'} \leq 1 ; ij, i'j', i''j'' \in E \& ij \neq i'j' \neq i''j'' \& D_C - d_{ii'} > 0 \& D_C - d_{jj''} > 0 \& l_{ij} + d_{ii'} + d_{jj''} > 2D_C \quad (30)$$

This cut is illustrated by the example provided in Fig. 11. Any located facility along edge  $ij$  cannot cover edges  $i'j'$  and  $i''j''$  simultaneously since  $(D_C - d_{ii'}) + (D_C - d_{jj''}) < l_{ij}$ .



**Fig. 11.** Impossible joint coverage

The subproblem (*SP*) of problem *P* is as follows.

*SP*:

$$\max 0 \tag{31}$$

$$y_{ij} + y_{ji} = l_{ij}; \quad ij \in E \mid i < j \tag{32}$$

$$c_{i'j'} + y_{ij} \leq D_C - d_{i'} + M(1 - \bar{\delta}_{ij'i'}) ; \quad i'j', ij \in E \ \& \ ij \neq i'j' \tag{33}$$

$$c_{i'j'} \leq l_{i'j'} \sum_{ij \in E \mid ij \neq i'j'} \bar{\delta}_{ij'i'} ; \quad i'j' \in E \tag{34}$$

$$e_{ij} \leq D_C \bar{x}_{ij} ; \quad ij \in E \mid i < j \tag{35}$$

$$e_{ij} \leq D_C \bar{x}_{ji} ; \quad ij \in E \mid i > j \tag{36}$$

$$e_{ij} + c_{ij} - y_{ij} \leq 0 ; \quad ij \in E \tag{37}$$

$$c_{ij} + c_{ji} + e_{ij} + e_{ji} \geq l_{ij} ; \quad ij \in E \mid i < j \tag{38}$$

$$y_{ij}, e_{ij}, c_{ij} \geq 0 ; \quad ij \in E \tag{39}$$

It is worth noting that the objective function could be minimization. However, since most of the constraints are of the form " $\leq$ ", it is more convenient to consider the maximization. Let  $\omega_{ij}$ ,  $\lambda_{ij'i'}$ ,  $\gamma_{ij}$ ,  $\theta_{ij}$ ,  $\alpha_{ij}$ , and  $\beta_{ij}$  denote the dual variables of constraints (32), (33), (34), (35-36), (37), and (38), respectively. Then, the dual subproblem (*DSP*) is as follows.

*DSP*:

$$\begin{aligned} \min \sum_{ij \in E \mid i < j} l_{ij} (\omega'_{ij} - \omega''_{ij} - \beta_{ij}) + \sum_{i'j', ij \in E \ \& \ ij \neq i'j'} (D_C - d_{i'} + M(1 - \bar{\delta}_{ij'i'})) \lambda_{ij'i'} + \sum_{i'j' \in E} \left( l_{i'j'} \sum_{ij \in E \mid ij \neq i'j'} \bar{\delta}_{ij'i'} \right) \gamma_{i'j'} \\ + \sum_{ij \in E \mid i < j} (D_C \bar{x}_{ij}) \theta_{ij} + \sum_{ij \in E \mid i > j} (D_C \bar{x}_{ji}) \theta_{ij} \end{aligned} \tag{40}$$

$$\omega'_{ij} - \omega''_{ij} + \sum_{i'j' \in E \ \& \ ij \neq i'j'} \lambda_{ij'i'} - \alpha_{ij} \geq 0 ; \quad ij \in E \mid i < j \tag{41}$$

$$\omega'_{ji} - \omega''_{ji} + \sum_{i'j' \in E \ \& \ ij \neq i'j'} \lambda_{ij'i'} - \alpha_{ij} \geq 0 ; \quad ij \in E \mid i > j \tag{42}$$

$$\sum_{ij \in E \ \& \ ij \neq i'j'} \lambda_{ij'i'} + \gamma_{i'j'} + \alpha_{i'j'} - \beta_{i'j'} \geq 0 ; \quad i'j' \in E \mid i' < j' \tag{43}$$

$$\sum_{ij \in E \ \& \ ij \neq i'j'} \lambda_{ij'i'} + \gamma_{i'j'} + \alpha_{i'j'} - \beta_{i'j'} \geq 0 ; \quad i'j' \in E \mid i' > j' \tag{44}$$

$$\theta_{ij} + \alpha_{ij} - \beta_{ij} \geq 0 ; \quad ij \in E \mid i < j \tag{45}$$

$$\theta_{ij} + \alpha_{ij} - \beta_{ji} \geq 0; \quad ij \in E \mid i > j \quad (46)$$

$$\lambda_{ijj'} \geq 0; \quad ij, i'j' \in E \mid i'j' \neq ij \quad (47)$$

$$\alpha_{ij}, \theta_{ij}, \gamma_{ij} \geq 0; \quad ij \in E; \quad \beta_{ij}, \omega'_{ij}, \omega''_{ij} \geq 0; \quad ij \in E \mid i < j \quad (48)$$

where the free variable  $\omega_{ij}$  is replaced by  $\omega'_{ij} - \omega''_{ij}$  and the bar sign over the *MP* variables denote their values. *DSP* always has a feasible region since it contains at least the zero vector. Thus, two cases would be occurred. If *DSP* is bounded, the lower bound and upper bound in the benders algorithm get equal and the algorithm is terminated. On the contrary, if the *DSP* is unbounded, the *SP* is infeasible and a feasibility cut must be added to the *MP*. Constraints of *DSP* constitute a polyhedral cone. Hence, every feasible solution (except for the origin) is a ray (Bertsimas & Tsitsiklis, 1997). *DSP* is unbounded if and only if, there is an extreme ray where the objective function decreases along it (Bertsimas & Tsitsiklis, 1997). Hence, the unboundedness of the *DSP* can be detected by solving the following modified *DSP* problem (*MDSP*).

*MDSP*:

$$\begin{aligned} & \min 0 \quad (49) \\ & \text{subject to (41-48) and} \end{aligned}$$

$$\begin{aligned} & \sum_{ij \in E \mid i < j} l_{ij} (\omega'_{ij} - \omega''_{ij} - \beta_{ij}) + \sum_{i'j', ij \in E \& ij \neq i'j'} (D_C - d_{i'j'} + M(1 - \bar{\delta}_{ijj'})) \lambda_{ijj'} + \sum_{i'j' \in E} \left( l_{i'j'} \sum_{ij \in E \mid ij \neq i'j'} \bar{\delta}_{ijj'} \right) \gamma_{i'j'} \\ & + \sum_{ij \in E \mid i < j} (D_C \bar{x}_{ij}) \theta_{ij} + \sum_{ij \in E \mid i > j} (D_C \bar{x}_{ji}) \theta_{ij} = -1 \end{aligned} \quad (50)$$

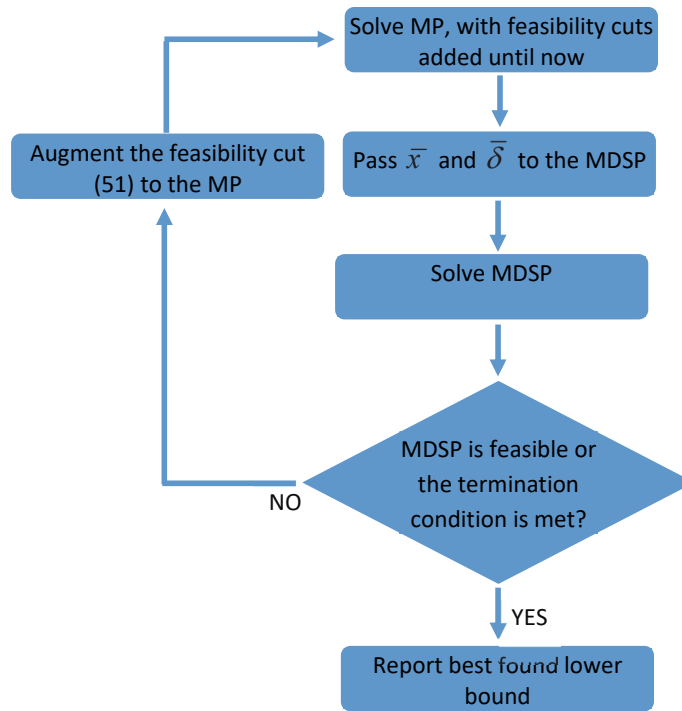
If the *MDSP* is feasible, the *DSP* is unbounded. In this case, cut (51) is added to the *MP*.

$$\begin{aligned} & \sum_{ij \in E \mid i < j} l_{ij} (\bar{\omega}'_{ij} - \bar{\omega}''_{ij} - \bar{\beta}_{ij}) + \sum_{i'j', ij \in E \& ij \neq i'j'} (D_C - d_{i'j'} + M(1 - \delta_{ijj'})) \bar{\lambda}_{ijj'} + \sum_{i'j' \in E} \left( l_{i'j'} \sum_{ij \in E \mid ij \neq i'j'} \delta_{ijj'} \right) \bar{\gamma}_{i'j'} \\ & + \sum_{ij \in E \mid i < j} (D_C x_{ij}) \bar{\theta}_{ij} + \sum_{ij \in E \mid i > j} (D_C x_{ji}) \bar{\theta}_{ij} \geq 0 \end{aligned} \quad (51)$$

where the bar sign over the dual variables denote their values. The summary of the Benders algorithm is shown in Fig. 12.

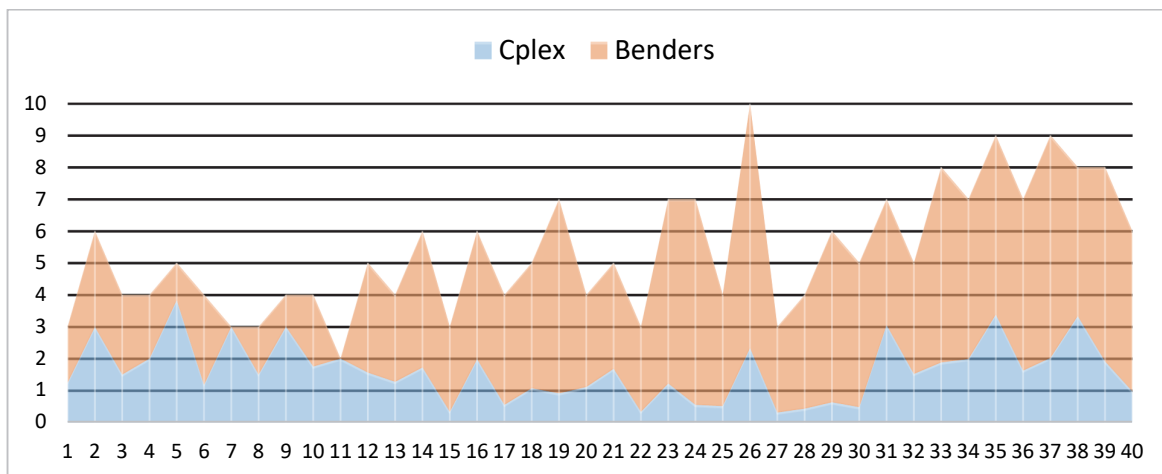
#### 4. Numerical results

We developed a heuristic algorithm for this problem. We evaluate the quality of the heuristic solutions by comparing them with the best found solutions and lower bounds generated by *CPLEX* and Benders algorithm. The heuristic solutions must be attained in reasonable time. The term *reasonable* depends on the nature of the problem (Bramel & Simchi-Levi, 1997). Real time decisions must be made in very short time. On the contrary, the process of decision making for long term problems can take a longer time. Facility location belongs to the class of long term decisions (Daskin, 2011). Hence, we set the time limit of 1200 seconds for the heuristic algorithm, the Benders algorithms, and *CPLEX*. The numerical experiments were run on a computer with 4.02 GHz Intel Core i7 processor and 14 GB RAM. The mathematical models were coded in GAMS IDE, version 24.1.2, and solved using the solver *CPLEX*, version 12.5.1. We generate random planar connected graphs as our test bed. To this end,  $N$  random points are generated in two-dimensional Euclidean plane. Then, Matlab is used to generate the Delaunay triangulation network, which implicitly constructs a connected planar graph. The length of each edge is calculated using the coordinates of its endpoints.  $D_C$  is randomly generated in the interval  $[0.5 \times \max\{l_{ij}\}, 1.5 \times \max\{l_{ij}\}]$ . Finally, the shortest path between each pair of vertices is calculated using Dijkstra's algorithm (Cormen et al., 2009).



**Fig. 12.** The Benders algorithm

Table 1 represents the results of solving 40 randomly generated problem instances. All run times are in seconds. Although the time limit for the *EDA* algorithm is 1000 seconds, in most of the instances it finds its best found solution in less than a few seconds. *LB* implies for the best found lower bound and *BF* denotes the best found solution.



**Fig. 13.** Lower bounds of Benders and CPLEX

Fig. 13 reports the difference between the lower bounds found by applying *CPLEX* on problem *P* and the lower bound found by the Benders algorithm. In all of the instances, the Benders algorithm finds tighter lower bounds than *CPLEX*, except for two instances in which both lower bounds are equal. This figure shows the superiority of Benders lower bounds. In 85 percent of the instances, Benders lower bounds are at least two times the lower bounds of *CPLEX*.

**Table 1**  
The detailed numerical results

<i>n</i>	#Instance	Benders		CPLEX			EDA	
		<i>LB</i>	Run time	<i>LB</i>	<i>BF</i>	Run time	<i>BF</i>	Run time
45	1	3	616	1.24	8	1202	4	3
	2	6	1202	2.99	11	1202	7	1124
	3	4	1202	1.5	9	1202	5	1
	4	4	1201	2	5	1201	4	1
	5	5	1201	3.83	7	1201	6	528
50	6	4	1203	1.17	7	1201	5	3
	7	3	1202	3	3	1202	3	4
	8	3	1201	1.5	4	1201	3	10
	9	4	1205	3	6	1201	5	80
	10	4	1201	1.75	9	1201	5	8
55	11	2	1047	2	2	462	2	7
	12	5	1202	1.56	14	1201	6	16
	13	4	1202	1.27	15	1201	5	6
	14	6	1202	1.72	64	1201	7	24
	15	3	1203	0.33	21	1138	3	6
60	16	6	1202	1.98	65	1201	7	2
	17	4	1203	0.55	43	1140	4	18
	18	5	1203	1.07	64	1201	6	19
	19	7	1203	0.91	45	1201	7	20
	20	4	1203	1.11	46	1201	5	66
65	21	5	1203	1.68	53	1201	5	6
	22	3	1203	0.33	27	1202	3	4
	23	7	1202	1.22	54	1201	7	13
	24	7	1203	0.55	65	1202	7	50
	25	4	1203	0.51	49	1202	5	137
70	26	10	1203	2.33	81	1202	12	1139
	27	3	1205	0.31	30	1202	3	364
	28	4	1137	0.43	59	1202	4	166
	29	6	1205	0.63	70	1202	6	5
	30	5	1204	0.48	58	1203	5	47
75	31	7	1203	3.05	17	1202	7	441
	32	5	1204	1.52	56	1202	5	124
	33	8	1203	1.88	59	1202	8	450
	34	7	1204	1.99	77	1203	7	455
	35	9	1203	3.38	23	1202	9	147
80	36	7	1205	1.62	64	1204	7	1021
	37	9	1204	2.02	68	1203	10	1200
	38	8	1203	3.33	60	1203	8	37
	39	8	1205	1.91	81	1203	8	383
	40	6	1205	0.97	46	1203	6	178

Fig. 14 compares the solutions of the *EDA* algorithm with the best found solution obtained through solving *P* by *CPLEX*. In other words, this figure depicts the savings in number of required facilities when the problem is solved by the *EDA* rather than solving problem *P* by *CPLEX*. On average, the number of used facilities by *CPLEX* is 6.5 times the ones used by the *EDA* algorithm. Further, in three-quarters of the instances, solving the problem by *EDA*, rather than solving *P* by *CPLEX*, results in at least 58% reduction in the number of required facilities. Fig. 15 represents the quality of the heuristic solutions by comparing them with the greatest lower bound found by *CPLEX* and Benders algorithm. On average, the number of used facilities by *EDA* algorithm are 0.08 larger than the maximum of two lower bounds. In 75% of the instances, *EDA* solutions match the maximum lower bound and hence are optimal. There is only one instance (i.e. #26) that *EDA* algorithm uses two additional facilities than the maximum lower bound.

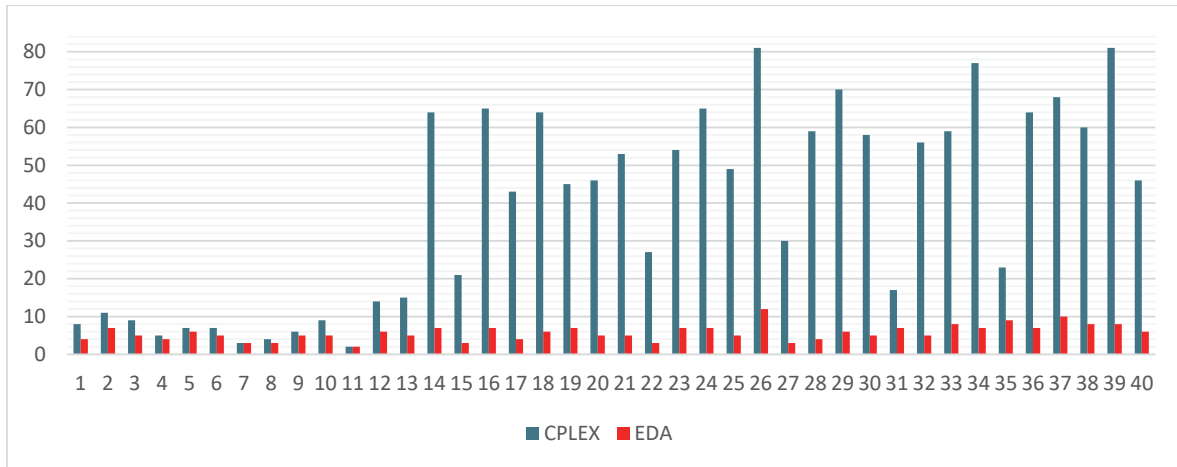


Fig. 14. Comparing EDA and CPLEX solutions (number of used facilities)

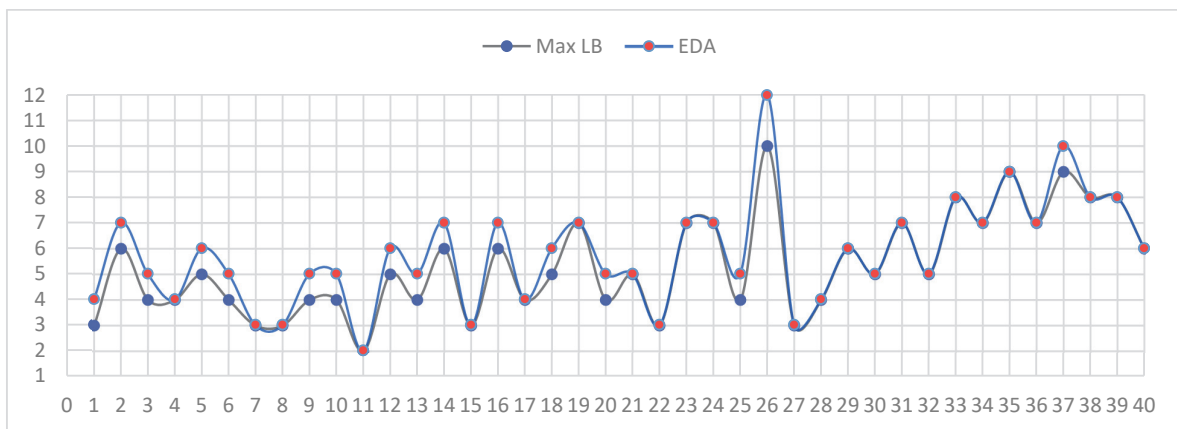


Fig. 15. EDA optimality gap

## 5. Concluding results and future research directions

This paper addressed the total edge covering problem in which the facilities are allowed to be located anywhere along the network. The mathematical programming formulation of the problem is developed for the first time. A matheuristic which iteratively solves the facility location on real and dummy nodes is proposed. In order to evaluate the quality of the heuristic solutions, a Benders algorithm is developed which generates tight lower bounds. Numerical results show the efficiency and effectiveness of the matheuristic algorithm for problems with 45 to 80 nodes.

This study addressed one of the facility location problems with continuous demand and location along the network. Many other facility location problems, like dispersion and center problems, can be attacked regarding the continuous location as well as the distributed demand along the edges. We considered uncapacitated facilities. The study can be extended regarding the capacity of the facilities and the density of the demand along the edges. For the ease of formulation, we assumed that at most one facility can be located along each edge. This assumption can be relaxed by allowing multiple facilities to be located along each edge.

## References

- Berman, O., Kalcsics, J., Krass, D.J.C. & Research, O. (2016). On covering location problems on networks with edge demand. *Computers & Operations Research*, 74, 214-227.
- Bertsimas, D. & Tsitsiklis, J.N. (1997). Introduction to linear optimization (Vol. 6): Athena Scientific Belmont, MA.

- Bramel, J. & Simchi-Levi, D. (1997). The logic of logistics: theory, algorithms, and applications for logistics management (pp. 175-240). New York: Springer.
- Cai, S., Li, Y., Hou, W. & Wang, H.J.I.S. (2019). Towards faster local search for minimum weight vertex cover on massive graphs. *Information Sciences*, 471, 64-79.
- Camm, J.D., Raturi, A.S. & Tsubakitani, S.J.I. (1990). Cutting big M down to size. *Interfaces*, 20(5), 61-66.
- Church, R.L. & Meadows, M.E.J.G.A. (1979). Location modeling utilizing maximum service distance criteria. *Geographical Analysis*, 11(4), 358-373.
- Church, R.L. & Murray, A.J.A.i.S.S. (2018). Location Covering Models. *Advances in Spatial Science*.
- Conrow, L., Murray, A.T. & Fischer, H.A.J.J.o.t.g. (2018). An optimization approach for equitable bicycle share station siting. *Journal of Transport Geography*, 69, 163-170.
- Cormen, T.H., Leiserson, C.E., Rivest, R.L. & Stein, C. (2009). Introduction to algorithms: MIT press.
- Daskin, M.S. (2011). *Network and discrete location: models, algorithms, and applications*. John Wiley & Sons.
- De Los Mozos, M.C.L. & Mesa, J.A.J.R.-O.R. (2000). The variance location problem on a network with continuously distributed demand. *RAIRO-Operations Research*, 34(2), 155-182.
- Farahani, R.Z., Asgari, N., Heidari, N., Hosseini, M., Goh, M.J.C. & Engineering, I. (2012). Covering problems in facility location: A review. *Computers & Industrial Engineering*, 62(1), 368-407.
- Groß, D.R.P., Hamacher, H.W., Horn, S. & Schöbel, A.J.T. (2009). Stop location design in public transportation networks: covering and accessibility objectives. *Top*, 17(2), 335.
- Hakimi, S.L.J.O.r. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, 13(3), 462-475.
- Heragu, S.S. (2008). Facilities design: Crc Press.
- Jayalakshmi, B., Singh, A.J.I.J.o.M.L. & Cybernetics. (2017). A hybrid artificial bee colony algorithm for the cooperative maximum covering location problem. *International Journal of Machine Learning and Cybernetics*, 8(2), 691-697.
- Laporte, G., Nickel, S. & da Gama, F.S. (2015). Location science (Vol. 528): Springer.
- Li, R., Hu, S., Zhang, H. & Yin, M.J.I.S. (2016). An efficient local search framework for the minimum weighted vertex cover problem. *Information Sciences*, 372, 428-445.
- Oğuz, M., Bektaş, T. & Bennell, J.A.J.E.J.o.O.R. (2018). Multicommodity flows and Benders decomposition for restricted continuous location problems. *European Journal of Operational Research*, 266(3), 851-863.
- Pandey, P. & Punnen, A.P.J.D.O. (2018). The generalized vertex cover problem and some variations. *Discrete Optimization*, 30, 121-143.
- ReVelle, C., Toregas, C. & Falkson, L.J.G.a. (1976). Applications of the location set-covering problem. *Geographical Analysis*, 8(1), 65-76.
- Sadigh, A.N., Mozafari, M. & Kashan, A.H.J.A.i.E.S. (2010). A mixed integer linear program and tabu search approach for the complementary edge covering problem. *Advances in Engineering Software*, 41(5), 762-768.
- Schobel, A., Hamacher, H.W., Liebers, A. & Wagner, D.J.A.-P.J.o.O.R. (2009). The continuous stop location problem in public transportation networks. *Asia-Pacific Journal of Operational Research*, 26(01), 13-30.
- Snyder, L.V. (2011). Covering problems. In Foundations of location analysis (pp. 109-135). Springer, Boston, MA.
- Taşkin, Z. C. (2010). Benders decomposition. *Wiley Encyclopedia of Operations Research and Management Science*.
- Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations research*, 19(6), 1363-1373.
- Wei, R., Murray, A. T., & Batta, R. (2014). A bounding-based solution approach for the continuous arc covering problem. *Journal of Geographical Systems*, 16(2), 161-182.

