Sensitivity analysis for p -median problems

Ágnes Vida
Institute of Informatics
University of Szeged
Vida.Agnes.1@stud.u-szeged.hu

Boglárka G.-Tóth
Institute of Informatics
University of Szeged
boglarka@inf.u-szeged.hu

ABSTRACT

Network location problems are commonly associated with real world situations such as locating a new facility in a city or setting up a new server in a computer network. In these real world situations changes can come up quite often, such as a closed road because of an accident, or a broken connection in the network. These kind of problems give the motivation of this paper. Our aim was to inspect, how an existing network operates, when something has changed and how sensitive a p -median solution is to the same changes. During our research we concentrated on deleting edges and solving p -median problem. In the p -median problem we try to locate p facilities so the sum of the distances between the facilities and the demand points is minimal. During the sensitivity analysis we deleted different amounts of the edges and we also tested the problem by locating 1, 2 or 3 medians to get a complex picture about the computational results. During our work, we concentrated on how the solution and the location of the medians change. To get a complex picture about the results we used two different graphs. After the tests we saw, that according to the shape of the graph, the number of the changes can be quite different. On the one hand, when we worked with a graph which is easy to cut, the changes of the solution was unstable, and relatively big, while with a well-balanced graph, the changes were not that significant. On the other hand, the location of the facilities did not change too much with either graph.

Categories and Subject Descriptors

G.2.2 [Mathematics of Computing]: Graph TheoryNetwork problems; G.1.6 [Mathematics of Computing]: Optimization

General Terms

Theory, application

Keywords

p -median, sensitivity analysis, Mixed Integer Programming

1. INTRODUCTION

Locating facilities is a common problem when we want to work with networks such as communication networks or traffic flows. In order to work with these networks in a mathematical aspect, we use an abstract model, a finite (often weighted, non-oriented) graph G which has a set of vertices (V) and connecting edges ($E \subseteq \{V, V\}$). Weights are commonly attached to the edges, which represent the cost of the transportation or the length of the road. Network locating problems include several different questions, such as absolute center and median problems or p -center and p -median problems. For a good introduction in facility location problems on networks, see [2].

In this paper we will concentrate on the p -median problem. The aim of the p -median problem is to find the optimal placement of p facilities on the network. The optimal location of the medians means that the sum of the distances between the vertices and the closest facility is minimal. To find the optimal location of the facilities, we must complete the following constraints: one demand point can be assigned to only one facility and maximum one facility can be placed at one vertex. We also have to provide that exactly p medians are located. To understand the p -median problem we studied the article by Hakimi [3].

During our research we always worked with non-oriented n -vertex graphs with edge weights, which can also be represented with a square matrix, called weight matrix. Since we are working on a graph, distance is always understood as shortest path on the graph. From the weight matrix we need to determine the distance matrix with some methods computing shortest paths on a graph.

The aim of this paper is to find out how sensitive a p -median solution is when we delete edges from the graph. The motivation is the use of traffic networks, to inspect how the traffic and the optimal locations of the p -medians change if something happens in the network. In this paper we concentrated on deleting edges with different likelihood, to inspect how the cost and the optimal location of the p -medians change. During our research we tested every case on two different graphs to get a more complex picture about the computational results.

Similar works has been done investigating the effect of changing network density in [4], and also studying facility reliability issues in Berman et al. [1].

Figure 1: Graph of Italy with cities on the islands (left), graph of Germany (right)

2. THE P -MEDIAN PROBLEM

The aim of solving a p -median problem is to locate exactly p medians by minimizing the sum of the distances between the demand points and the located facilities and also completing a few constraints mentioned below. In order to solve the problem we used the following model which is based on a non-oriented weighted n -vertex graph. We worked with the graph as a set of vertices (V) connected by weighted edges ($E \subseteq \{V, V\}$).

$$\min \sum_{i,j \in V} d_{ij} y_{ij} \quad (1)$$

$$\text{s.t.} \quad \sum_{j \in V} y_{ij} = 1 \quad \forall i \in V \quad (2)$$

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

$$y_{ij} \leq z_j \quad \forall i, j \in V; \quad (4)$$

To represent the distances between the vertices we used a parameter d_{ij} , which is a squared matrix. The j -th column of the i -th row represents the length of the shortest path between the i -th and j -th vertex.

We also used two set of variables: y_{ij} and z_i . The binary variable y_{ij} is to sign that the i -th demand point is supplied by the j -th facility and z_i is a binary variable to sign if we locate a facility on vertex i . The aim of the model is to minimize the sum of the distances between the demand points and the facilities. The optimal solution must complete three constraints: (2) gives the condition, that exactly one facility has to be assigned to one vertex; the number of located facilities must be p which is required by constraint (3); and finally, that a demand point can only be assigned to a facility, that is located, see (4).

Our aim is to minimize the sum of the distances between the demand points and the closest facility, but we do not have any direct constraint for that. That is because the model minimizes the sum of $d_{ij} y_{ij}$, and so in this way the model automatically chooses the best solution, which is the minimal distance between a demand point and the assigned facility.

3. METHODS

Our choice to get the computational results was AMPL, which is A Mathematical Programming Language. First we implemented the model to solve the p -median problem. In order to obtain the optimal solution of this Mixed Integer Programming (MIP) problem we used a commercial solver called CPLEX.

Our goal was to inspect how the solution reacts to changes, especially deleting edges. At first, we always ran the algorithm with no changes to inspect the original solution, than we modified it with different amounts. We increased the number of deleted edges always by 10%. After deleting 40% of the edges we decided to stop, because we noticed that the graph fell apart. We also checked how the solution operates when we want to locate 1, 2 or 3 medians with the different number of deleted edges. In every variation of the number of deleted edges and the number of located medians we made 20 test runs.

In order to delete edges, first we made a mapping, so we could work with the edges like an ordered set. Then we randomly choose the right number of edges to delete from the graph. To implement this random choose, we used some of the built-in functions of AMPL. Deleting the desired amount of edges was done uniformly. In a cycle we have generated a random integer value k from 1 to the number of still existing edges and removed the k th edge, until we removed the required number of edges.

After deleting the chosen edges, we determined the distance matrix on the modified graph. In order to get the distances, we used the definition of shortest path, and the Ford-Fulkerson algorithm.

4. RESULTS AND ANALYSIS

The algorithm was tested on two data files, modeling the largest cities and roads of Italy and Germany. In order to have a better view on the results, we draw the graphs using the real coordinates of the cities in Figure 1. We segmented the tests according to how many medians are located (1, 2 or 3) and the number of the deleted edges (0, 10%, 20%, 30% and 40%). We made 20 test runs for every segment and reported the average, the minimum and maximum changes. The tests show that the shape of the graph influences the results a lot.

The results for the comparison of the solutions in changes of the cost, as well as in the changes in the locations for both the Italian and German networks can be seen in Table 1. Next we discuss the results that can be read from the table.

4.1 Changes of the cost

In the Italian graph, when we deleted 10% of the edges, the solution increased by 309% on average. This is quite huge change as in the German graph this increase was only 10%, and all changes was smaller than 134%. But let us concentrate first the results for the Italian graph. We have found that the solution was very unstable for this graph. When we located 1 median, by deleting 10% of the graph the minimal increase was by 22% while the maximal increase was around 1086% and the average increase was 309%. By deleting the double amount of the edges, the average increase of the objective function was also about double (615%), the minimal and maximal increase was 82% and 1314%, respectively. We also tested the results during deleting 30% and 40% of the edges. With a 30% loss, the average increase of the objective value was 1275%. After deleting 40% percent of the edges we stopped increasing the number of deleted edges, because we found that the increase of the objective value is too high (1387% on average), and we have seen that this happened because the graph fell apart. We can see a big jump in the minimum changes in costs from 20% to 30%, when from 82% it grew to 737%. Similar, but not so evident jumps can be seen in the minimum change when 2 or 3 medians are located.

When we located two medians, the results were quite similar. The main difference showed, when we deleted only 10% of the edges. In this case the cost increased by only 51%. By deleting 20% of the edges the average increase of the cost was around 345%, by deleting 30% of the edges, the cost increased by 925% on average, and next to a 40% loss of the edges, the objective value increased by 1421%.

The problem of locating 3 medians showed similar results as the 2-median problem. The average increase of the cost was 79% when we only deleted 10% of the edges and 1352% when we deleted 40% of the edges. Altogether, we can see that locating more medians makes less changes in cost in general for this graph (except 40% deleted edges), although not significantly.

When we used the graph of Germany, the results were different. The tendency was the same, but the amount of the rise of the solution was much lower. First we located only one facility. By deleting 10% of the edges and locating one facility the cost increased by 10% (minimum 1% and maximum 23%). By deleting more edges, the solution increased slightly, but not as much as with the Italian graph. When we deleted 40% of the edges, the cost increased with 84% on average (minimum 57%, maximum 124%). As with the previous graph, we repeated the test by locating two and then three facilities. The results were not that different, by deleting 10% of the edges the increase of the cost was 10% with 2 medians and 9% with 3 medians. Even by deleting 40% of the edges the increase of the cost was around 117%. In this case, the changes compared for the number of medians is quite balanced, as opposed to the Italian case. During the tests we could see, that there are vertices where facilities are more likely to be placed.

4.2 Changes of the locations

We also paid attention to the location of the medians. We have made a ranking on the number of times a vertex was selected as a median in the 20 runs. When we had to locate only one facility with the Italian graph, we saw that the original location was only the 11th most likely place to put the facility at. With the German graph the result was much better: the original placement of the median was the 4th in the ranking. When we solved a 2-median problem, the results were the same with the two graphs: the original locations of the medians were in the 5 most likely vertices to put a facility at. With 3-median problem the results were also similar. The original location of the medians were still in the top 5 vertices to locate a facility on. All in all we can say, that the location of the facilities can change according to the number of deleted edges, but the original locations are almost always in the top 5 vertices.

During our research we also made computations to inspect, how far the new locations from the old ones are. To get an objective picture about the results we always used the original graphs and the shortest distances between the old and the new locations of the medians. When we located more than 1 median, we choose the smallest change between the old and the new locations in pairs. Although the values were higher with the German graph, the tendency was the same. With the Italian graph when we located only one facility and deleted 10% of the edges the distance between the old and the new one was 16.6 on average (minimum 4, maximum 51). Even with deleting 40% of the edges the average distance was 61.1 (minimum 8, maximum 81). This small amount of change is interesting if we see, that the cost increased much more under these conditions. When we located 2 medians, the changes were even smaller. With a 10% loss, the distance between the old and the new facilities was 3.9. When we deleted 40% of the edges the distance was still around 9%. When we located 3 medians, the new locations were farther than with the 2- and 3-median problems. In this case the minimal change was around 10 (by deleting 10% of the edges) and the maximal change was about 44 (by deleting 30% of the edges).

With the German graph we did not see significant differences. During 1-median problem and a 10% loss the dis-

Table 1: Results for the comparison of the solutions in changes of the cost, as well as in the changes in the locations for both the Italian and German networks.

	Number of facilities	Deleted edges	Change of the cost			Change of the location		
			Min	Avg	Max	Min	Avg	Max
Italy	1	10%	22%	309%	1086%	4	16.6	51
		20%	82%	615%	1314%	0	18.4	51
		30%	737%	1275%	1524%	4	39.6	62
		40%	835%	1387%	1613%	8	61.1	98
	2	10%	1%	51%	170%	0	3.9	39
		20%	106%	345%	718%	0	8.3	66
		30%	325%	925%	1546%	0	8.2	55
		40%	878%	1421%	1831%	0	9.0	47
	3	10%	2%	79%	150%	0	10.4	132
		20%	34%	233%	613%	0	31.7	128
		30%	298%	616%	907%	0	44.3	153
		40%	904%	1352%	1854%	0	40.0	124
Germany	1	10%	1%	10%	23%	0	81.6	247
		20%	9%	27%	49%	0	131.9	265
		30%	23%	57%	86%	100	166.8	264
		40%	57%	84%	124%	178	188.4	264
	2	10%	2%	10%	19%	0	40.3	195
		20%	13%	26%	47%	0	75.7	276
		30%	21%	59%	98%	0	105.7	276
		40%	81%	104%	127%	0	105.2	276
	3	10%	2%	9%	19%	0	69.9	471
		20%	4%	24%	42%	0	117.5	614
		30%	34%	55%	79%	0	84.7	276
		40%	94%	117%	134%	0	132.3	440

tance between the old and the new location of the facilities was 81.6 on average (minimum 0, maximum 247). Even with a 40% loss of the edges the distance was 188.35 on average (minimum 178, maximum 264). When we located 2 medians, the average change was around 40 (minimum 0, maximum 195) with deleting 10% of the edges and 105.2 (minimum 0, maximum 276) with a 40% loss. With the 3-median problem, the minimal change was about 70 and the maximal change was about 132.

5. CONCLUSION AND FUTURE PLANS

By the results we can see that the change of the solution during deleting edges highly depends on the shape of the graph we work with. If the graph has just a few connections between two main parts, deleting of the edges can cause very high costs. If we work with a well-balanced graph, even by deleting 40% of the edges the solution changes just a little. We also could see that the location of the facilities do not change too much. In the future we are planning to inspect how a non-oriented weighted graph reacts to p -median problem when we change the weights of edges, like when traffic changes or when we add new edges. We also would like to check if we can make clusters according to how likely that a facility is located in a vertex. Furthermore we want to inspect how a non-oriented weighted graph reacts to other problems such as p -center problem during similar conditions.

6. ACKNOWLEDGMENTS

This research was supported by the European Union and co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002).

7. REFERENCES

- [1] O. Berman, D. Krass, and M. B. C. Menezes. Facility reliability issues in network p -median problems: Strategic centralization and co-location effects. *Operations Research*, 55(2):332–350, 2007.
- [2] M. S. Daskin. *Network and Discrete Location: Models, Algorithms and Applications*. John Wiley and Sons, New York, 1995.
- [3] S. L. Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459, 1964.
- [4] X. Zhao, K. Carling, Z. Dan, and J. Håkansson. Network density and the p -median solution. Technical report, Högskolan Dalarna, Borlänge, 2013.