

Analysis of Facility Systems' Reliability Subject to Edge Failures: Based on the p -Median Problem

Zongtian Wei^{1,2}, Huayong Xiao¹, Yuxi Quan²

¹Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, China

²Department of Mathematics, Xi'an University of Architecture and Technology, Xi'an, China

E-mail: ztwei@xauat.edu.cn, huayongx@nwpu.edu.cn, quanyuxi1216@126.com

Received August 28, 2011; revised October 3, 2011; accepted October 17, 2011

Abstract

We view a facility system as a kind of supply chain and model it as a connected graph in which the nodes represent suppliers, distribution centers or customers and the edges represent the paths of goods or information. The efficiency, and hence the reliability, of a facility system is to a large degree adversely affected by the edge failures in the network. In this paper, we consider facility systems' reliability analysis based on the classical p -median problem when subject to edge failures. We formulate two models based on deterministic case and stochastic case to measure the loss in efficiency due to edge failures and give computational results and reliability envelopes for a specific example.

Keywords: Facility System, Reliability, Edge Failure, p -Median Problem, Operating Efficiency

1. Introduction

Church and Scaparra [1] defined critical infrastructure as those elements which are necessary for life line support and safety. They include such systems as communication systems, transportation systems, water and sewer systems, health services facilities, etc. Each of these systems has unique properties that may define specific issues in operation and management in order to provide a consistent and continuing level of operation.

Every facility system in operation maybe faces various disruptions. Such disruptions have begun to receive significant attention from practitioners and researchers after the terrorist attacks of September 11, 2001. One reason for this growing interest is the spate of major disruptions since the new century such as the foot-and-mouth disease scare in the UK in 2001, the terrorist attacks on September 11, 2001 and the west-coast port lockout in 2002 in the US, the Asia SARS outbreak in 2003, the Indian Ocean tsunami on December 26, 2004, and the Wenchuan earthquake on May 12, 2008 in China, etc. In fact, various natural disasters or intentional strikes constantly occur in the world, e.g., congestions, inclement weather, earthquakes, debris flows, sandstorms or terrorist attacks.

Another reason is that, firms are much less vertically integrated than they were in the past, and their supply chains are increasingly global. Today, firms tend to as-

semble final products from increasingly complex components, which are procured from suppliers rather than produced in-house. These suppliers are located throughout the globe, many in regions that are unstable politically or economically or subject to wars and natural disasters.

Facility system disruptions can have significant physical costs [2]. Therefore, Reliability, Robustness and Resilience (3R) are receiving high attention in the design of facility systems [3].

The reliability of a facility system is the probability that all suppliers are operable [4]. This concept is introduced from the network reliability theory. Generally speaking, the key difference between networks reliability and facility systems reliability is that the former are primarily concerned with connectivity; they consider the cost of constructing the network but not the cost that results from a disruption, whereas the latter consider both types of costs and generally assume connectivity after a disruption [5].

In location science, the problem of locating p -supply facilities that yields the smallest weighted distance is called the p -median problem (PMP). The PMP has been the subject of considerable research, starting with the theorems of Hakimi [6,7], the first heuristic of Teitz and Bart [8] and an integer linear programming model of Reville and Swain [9]. Church [10] provided a detailed

summary of different approaches for solving the PMP.

A number of papers in the location literature have addressed the problem of finding the optimal location of protection devices to reduce the impact of possible disruptions to infrastructure systems. For example, Carr *et al.* [11] presented a model for optimizing the placement of sensors in water supply networks to detect maliciously injected contaminants. James and Salhi [12] investigated the problem of placing protection devices in electrical supply networks to reduce the amount of outage time.

Church *et al.* [13] presented a model called the r -interdiction median problem. This model can be used to identify which r of the existing set of p -facilities, when interdicted or lost impacts delivery efficiency the most. Such a model can be used to identify the worst case of loss, when losing a pre-specified number of facilities. The model is restricted in two ways: it is based upon the assumption that the terrorist or interdictor is successful in each and every strike, and it is also based upon the assumption that exactly r facilities will be struck and lost. Such a model does address a worst case scenario, but it does not exactly capture the issues that would be key to understanding the range of failures and possible outcomes.

In [1], the authors argued that first, it is important to recognize that a strike or disaster may not impair a facility's operation. That is, a terrorist strike may be successful only a certain percentage of the time. The same is true for a natural disaster. When it does occur, there is a threat that operations at a facility may need to be suspended, but it is not absolute. Second, interdiction may not be intelligent when the strikes impact a non-critical facility. Although it is important to model "worst-case" scenarios, it is also important to model and understand the range of possible failures and impacts. Therefore, they proposed a family of models which can be used to model the range of possible impacts associated with the threat of losing one or more facilities to a natural disaster or intentional strike. They show how to model deterministic loss and probabilistic loss. In addition, they presented results associated with the application of the worst case and the best case expected loss models to a data set.

There is a mature literature on reliable, robust and resilient network (e.g., supply chain network (SCN)) design and analysis under component failures. Unfortunately, so far we have not found the explicit study of facility systems' reliability subject to edge failures. In fact, generally the reliability, and hence the efficiency, of a facility systems is to a large degree adversely affected by failures of the edges. Thus the network-based facility system reliability models to be investigated are more practical and closer to the reality of facility system man-

agement.

In this paper, adopting the facility location analysis framework, we will mainly consider facility systems' reliability analysis based on the classical p -median problem when subject to edge disruptions. The remainder of the paper is organized as follows. In Section 2, we formulate two reliability models based on the p -median problem and edge failures. We use a scenario based algorithm to compute a specific example and give the results and reliability envelopes in Section 3. Section 4 is a summary of this paper.

In the following, by "loss" we refer to the edge disruptions (failures) mentioned above or, sometimes the necessary closure.

2. Facility System Reliability Analysis Models

The PMP locates p facilities to minimize the demand-weighted total (or average) distance between demands and the nearest facility. The exact opposite of the p -median problem occurs when an existing system of p facilities, which may or may not be located optimally. When either closing or considering the loss of one or more transportation lines by a disaster, the basic question is what happens to the operating efficiency of the system.

2.1. The Deterministic Reliability Models

The following are the notations for our formulations.

Sets

I : set of customers, indexed by i .

J : set of potential facility locations, indexed by j .

Parameters

h_i : demand at customer $i \in I$.

We view a facility system as a weighted connected simple graph $G(V, E, H, D)$, where $V = I \cup J$; E is the edge set with the edges denoting goods or information paths; H is the vertex weight set with the weight h_i of vertex i , denoting the demand of customer i , and D is the edge weight set with the weight d_{ij} of edge (i, j) , denoting the length (e.g., distance) between i and j under the existing conditions. By d_{ij} we also denote the shortest path length (or distance) between i and j if $(i, j) \notin E$. Evidently, $0 \leq d_{ij} < \infty$.

Let C be the opened facility (server) set of an existing facility system G , and $F \subseteq E$ be the potential failure edge set, where the edge failure is defined as an edge losses its designed function completely. Therefore, a failed edge in G is equivalent to delete (or close) the corresponding line from the facility system. We also assume that the edge failures are independent and multiple edge failures may occur simultaneously.

Let S_r be the set of scenarios corresponding to the closure or deletion of r edges from G , *i.e.*, every $s \in S_r$ explicitly specifies the failed r edges in F . Let d_{ijs} be the shortest distance between customer i and facility j in scenario s . Define $N_{is} = \{j \in C : d_{ijs} < +\infty\}$. Assume that in any scenario $s \in S_r$, customer i is served by an opened facility $j \in C$ which is the nearest one from i if $N_{is} \neq \emptyset$. Associated with each customer $i (\in I)$ is a per-unit penalty cost θ_i that represents the cost of not serving the customer if $N_{is} = \emptyset$. θ_i may represent a lost-sales cost, or the cost to pay a competitor to serve the customer temporarily. Denote $Y_{ijs} = 1$, if customer i is assigned to facility j in scenario s ; 0, otherwise.

We formulate the deterministic reliability model (DRM) as the following integer-linear programming problem:

$$\min \left(\sum_{s \in S_r} \sum_{i \in I, N_{is} \neq \emptyset} \sum_{j \in C} h_i d_{ijs} Y_{ijs} + \sum_{i \in I, N_{is} = \emptyset} h_i \theta_i \right) \quad (1)$$

$$\text{s.t. } \sum_{j \in C} Y_{ijs} \leq 1, \forall i \in I, s \in S_r$$

$$Y_{ijs} \in \{0, 1\}, \forall i \in I, j \in C, s \in S_r \quad (2)$$

The objective function selects r failed edges from F in order to minimize the resulting weighted distance. Constraints (1) require that each customer be served by at most one server in any scenario. Constraints (2) require the assignment variables to be binary.

For a given edge loss level r (the number of closed or deleted edges), this model can be used to evaluate a facility system's operational efficiency under the best case, namely the minimal loss of the system's efficiency. Changing the "min" to "max" in the objective function, then we obtain the worst case model, that is the model to measure the maximal increase in weighted distance under the edge failure level r .

2.2. The Stochastic Reliability Models

The reliability model formulated above is based upon a deterministic analysis. Up to this point we have modeled edge loss a certainty. We now consider the case where loss is not a certainty upon an edge failure. Usually, the chances of losing an edge are based upon some probability. We wish to derive the maximal or minimal expected efficiencies associated with an existing system. To do so we need to identify both the worst case and the best case expected outcomes.

Let $F \subseteq E$ be the target edge (the potential failure edge) set of an attack. Assume that an attacker can hit each edge in F at most once and that the edges in F will be hit simultaneously.

Let S_r be the scenario set when r edges in F

have been attacked. Each $s \in S_r$ specifies which r edges in F have been attacked. Denote

$E_s = \{e \in F | e \text{ is attacked in scenarios}\}$, then any

$F_s \subseteq E_s$ can be used to represent a failed edge set in scenario s , so we call F_s the sub-scenario of s . Let

d_{ijF_s} be the shortest distance between customer i and an opened facility j in scenario F_s . Define

$N_{iF_s} = \{j \in C : d_{ijF_s} < +\infty\}$. Assume that in any scenario F_s , customer i is served by an opened facility $j \in C$

which is the nearest one from i if $N_{iF_s} \neq \emptyset$. Let p_j be the failure probability of edge $j \in F$ after one attack.

It is easy to see that scenario F_s occurs with probability

$$p_{F_s} = \prod_{j \in F_s} p_j \prod_{j \in E_s \setminus F_s} (1 - p_j).$$

Denote the assignment variables as $Y_{ijF_s} = 1$, if customer i is served by facility j in scenario F_s ; 0, otherwise.

Let C be the opened facility (server) set of an existing facility system. We formulate the stochastic reliability model (SRM) as the following integer-linear programming problem (The penalty cost θ_i are defined in Section 3.1.):

$$\min \left(\sum_{s \in S_r} p_{F_s} \sum_{F_s \subseteq E_s} \sum_{i \in I, N_{iF_s} \neq \emptyset} \sum_{j \in C} h_i d_{ijF_s} Y_{ijF_s} + \sum_{i \in I, N_{iF_s} = \emptyset} h_i \theta_i \right) \quad (3)$$

$$\text{s.t. } \sum_{j \in C} Y_{ijF_s} \leq 1, \forall i \in I, F_s \subseteq E_s$$

$$Y_{ijF_s} \in \{0, 1\}, \forall i \in I, j \in C, F_s \subseteq E_s \quad (4)$$

The objective function selects r edges from F to minimize the weighted distance expectation after r edges in F have been attacked. Constraints (3) require that each customer be served by at most one server in any scenario F_s . Constraints (4) require the assignment variables to be binary.

For a given failure level r , this model can be used to evaluate a facility system's operational efficiency under the best case. Changing "min" in the objective function to "max", then we obtain the worst case model, *i.e.*, the model to measure the maximal increase in expected weighted distance under failure level r .

3. The Reliability Envelopes

The models described above can be applied to a given facility system over a range of edge loss level r . We use the weighted distance to measure the efficiency of a facility system and efficiency is measured at 100% if all edges are operating. If an edge is lost due to a natural disaster, intentional strike or planned closure, then the efficiency is lost and overall efficiency decreases. If many edges exist, then there exist several possible out-

comes of losing just one edge. One can easily enumerate each of the possible ways of losing one edge as well as calculate the impact of each possible loss in terms of changes in efficiency. The results of this series of calculations will define a range of losses from the best case (*i.e.* the least decrease in efficiency) to the worst case (*i.e.* the greatest decrease in efficiency). We then have a region defined by an upper curve and a lower curve, where the upper and the lower curve represent the solutions of the least or the greatest impact associated with a given loss level, respectively. The region depicted between these two curves can be defined as the operational envelope or reliability envelope. For a given edge loss level, this envelope specifies the range of possible system performance from the best-case to the worst-case. Actual performance will fall within this range.

In this section we apply the DRM and the SRM to a data set to generate reliability envelopes. Our data set is derived from the 2008 China census data: a 49-node set consisting of the capitals of all the provinces in China plus the two special administrative regions Hong Kong and Macau, as well as other 15 big cities. The demand of city i , h_i is settled to be the city's administrative region population divided by 10^4 . The transportation links are the recent national highways and the transportation cost between city i and city j , d_{ij} equal to the distance between the two cities.

We optimally solve a 6-median problem in order to obtain an existing facility system.

Figure 1 shows the optimal solution, where the 6 distribution centers are city 1, city 7, city 10, city 19, city 22 and city 44, and the edges marked by red color represent the delivery routes from each distribution center to its customers.

Given this operating system of 6 facilities and a potential failure edges set

$$F = \{1(5, 47), 2(1, 3), 3(7, 40), 4(10, 38), 5(22, 23), 6(22, 25), 7(27, 44), 8(19, 21)\},$$

which is consisted of 8 edges (the sub-graph $G - F$ is connected, *i.e.*, both of N_{is} and N_{if_s} are not ϕ), we then consider edge loss level from 1 to 8 and solve the worst case DRM and the best case DRM. The solutions are given in **Table 1**.

In **Table 1**, for each edge loss level, the objective function values and efficiency for each case are also given as a percentage, where 100% represents the operating level before edge failures. Of cause, the edges shown in volume 6 of **Table 1** are the most important objective of protection.

Figure 2 presents the values of operational efficiencies (in percent) as a graph, depicting the lower and the upper boundaries of the reliability envelope. Notice that the greatest marginal impact for the worst case occurs when the edge loss level is small while for the best case occurs when the edge loss level is great.

It is also important to note that the greatest difference

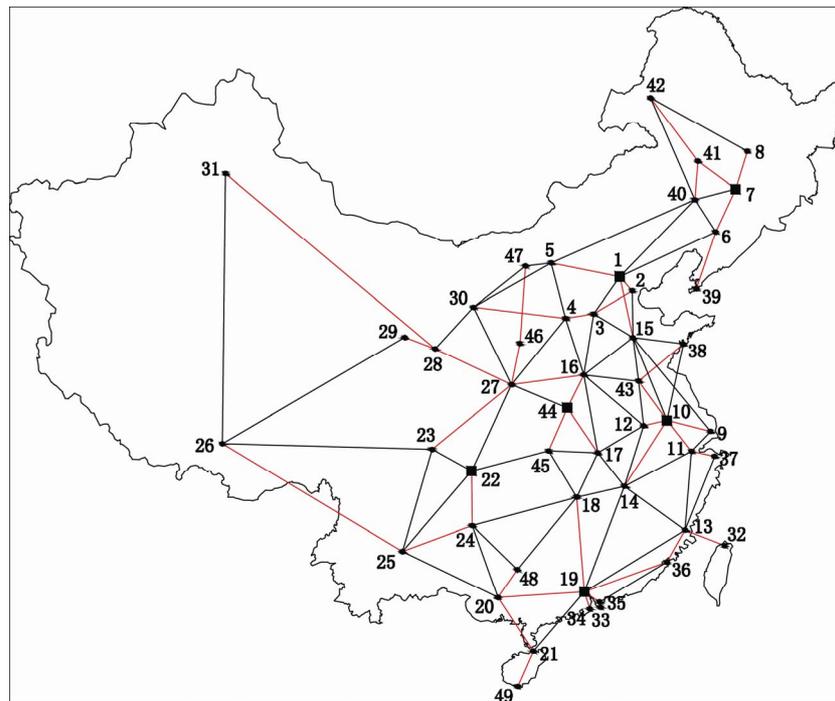


Figure 1. Optimal solution of a PMP with $p = 6$.

Table 1. Results of the worst-case DRM and the best-case DRM.

Level	Best-Case			Worst-Case		
<i>r</i>	Objec. Value	Failed Edges	Efficiency	Objec. Value	Failed Edges	Efficiency
0	11,572,831	-	100%	11,572,831	-	100%
1	11,599,381	6	99.77%	12,989,073	5	89.10%
2	11,669,755	3, 6	99.17%	13,558,323	5, 7	83.54%
3	11,742,050	3, 4, 6	98.56%	13,852,403	2, 5, 7	83.54%
4	11,887,640	3, 4, 6, 8	97.35%	14,059,421	1, 2, 5, 7	82.31%
5	12,051,692	1, 3, 4, 6, 8	96.03%	14,205,011	1, 2, 5, 7, 8	81.47%
6	12,345,772	1, 2, 3, 4, 6, 8	93.74%	14,277,306	1, 2, 4, 5, 7, 8	81.06%
7	12,735,040	1, 2, 3, 4, 6, 7, 8	90.87%	14,347,680	1, 2, 3, 4, 5, 7, 8	80.66%
8	14,376,780	1, 2, 3, 4, 5, 6, 7, 8	80.50%	14,376,780	1, 2, 3, 4, 5, 6, 7, 8	80.50%

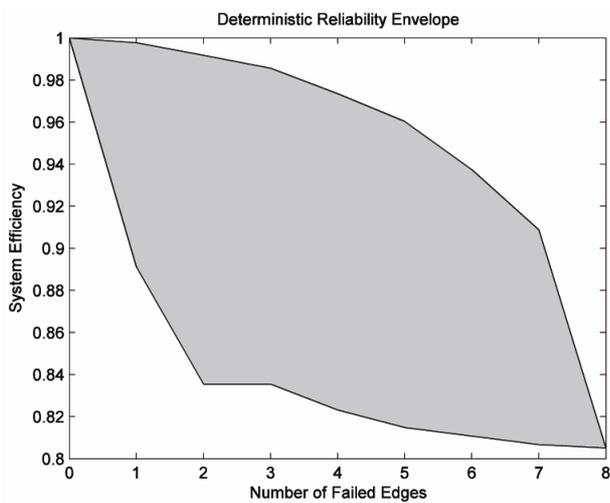


Figure 2. Reliability envelope of solutions in Table 1.

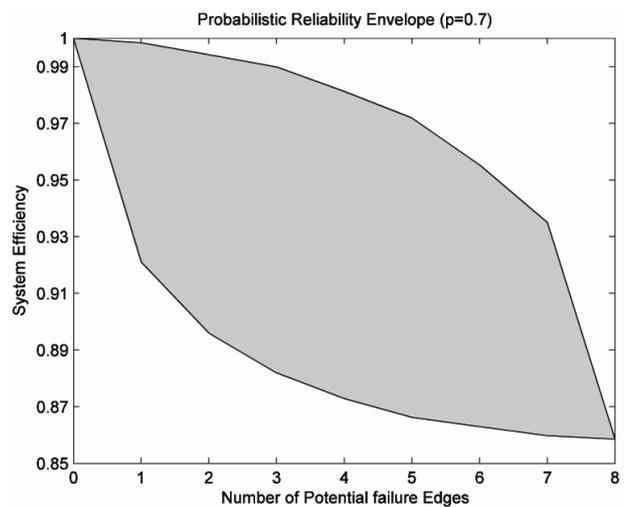


Figure 3. Reliability envelope of solutions in Table 2.

between the worst case and the best case, or the thickness of the envelop occurs when the edge loss level is moderate.

By using the same data set as in the above DRM, we then solve the worst case and the best case SRM with edge failure probability $p = 0.3$ and $p = 0.7$, respectively. The solutions of the latter and the corresponding reliabil-

ity envelope are showed in **Table 2** and **Figure 3**, respectively. Notice that the characteristics of this reliability envelope are similar to that of the DRM except that, on the same edge loss level, the efficiency losses of the SRM are less than that of the DRM, since we assume that the failure probability of an edge under a strike is 1 in the DRM. We can also observe from **Table 2** that, for

Table 2. Results of the worst case SRM and the best case SRM with edge failure probability 0.7.

Level	Best-Case			Worst-Case		
<i>r</i>	Objec. Value	Attacked Edges	Efficiency	Objec. Value	Attacked Edges	Efficiency
0	11,572,831	-	100%	11,572,831	-	100%
1	11,591,416	6	99.84%	12,564,200	5	92.11%
2	11,640,678	3, 6	99.42%	12,915,856	5, 7	89.60%
3	11,691,284	3, 4, 6	98.99%	13,121,712	2, 5, 7	88.20%
4	11,793,197	3, 4, 6, 8	98.13%	13,257,602	1, 2, 5, 7	87.29%
5	11,908,034	1, 3, 4, 6, 8	97.19%	13,359,515	1, 2, 5, 7, 8	86.63%
6	12,113,890	1, 2, 3, 4, 6, 8	95.53%	13,410,122	1, 2, 4, 5, 7, 8	86.30%
7	12,377,354	1, 2, 3, 4, 6, 7, 8	93.50%	13,459,383	1, 2, 3, 4, 5, 7, 8	85.98%
8	13,479,218	1, 2, 3, 4, 5, 6, 7, 8	85.86%	13,479,218	1, 2, 3, 4, 5, 6, 7, 8	85.86%

a given edge attacked level, the most important edges of protection are shown in volume 6.

4. Summary and Conclusions

In this paper, we propose two types of scenario based facility location models in order to analyze the reliability of an existing facility system when subject to edge failures. We distinguish deterministic and stochastic cases to formulate and compute a specific example. Reliability envelopes in these two different cases are also given. The information in the reliability envelopes can be very useful in looking at ways to protect a facility system. Whether the protection is against a natural disaster or intentional strike, reducing the probability of success even by modest amounts could have an impact on system efficiency. For example, this could be done by placing extra strength in key sections such as bridges or tunnels which spaced in disaster-prone areas, or by adding a surveillance system with guards to help protect against an intruder. Either techniques may not completely eliminate a loss, by reduce the edge failure probability to zero, but such strategies may generate more benefits in terms of improved expected system operating efficiencies than what it might cost. Therefore, the value of our analysis could lead to higher levels of safety as well as efficient levels of resource allocation for security measures (whether that involves a possible natural disaster or an attacker).

While a large body of literature focus on the reliable and robust facility system design and analysis under component failures, the existing works are mainly concentrated on the "node" (e.g., suppliers, distribution centers) losses. To the best of our knowledge, researchers and practitioners have not paid enough attention to the impact of edge failures to facility systems' efficiency by so far. Comparing to the related works done in this field, our work have at least the following innovations.

Firstly, combining edge failures into facility systems reliability analysis is more realistic than only considering vertex failures. In fact, natural disasters or intentional attacks damage the edges of a facility system more easily. Secondly, in the PMP and other uncapacitated facility location problems, when one or more "nodes" have failed, the overall supplement of the facility system will decrease dramatically but the total demand does not change. If in this situation all demands must to be met, then every node must has no any capacity limit. However, the capacity of nodes are designed a priori, when a vertex failure happen, how can it's adjacent nodes guarantee the increased demands, let alone more than one vertex failure occur simultaneously?

The recovery time for a failed edge maybe shorter than

that for a failed vertex, but this is not always the case. So we do not explicitly point out the time horizon in our models. In addition, the evaluation of edge failure probability is important and difficult, and we will discuss this question in another work. We set the failure probability of all edges as the same when solving the example since we aimed to demonstrate the impact of edge failures to the efficiency of a facility system.

5. Acknowledgements

This paper was supported by the Open Fund of Xi'an Jiaotong University (No. 2010-4), SXESF (No. 09JK545) and BSF (No. JC0924).

6. References

- [1] R. L. Church and M. P. Scaparra, "Critical Infrastructure," Springer, Berlin, 2007.
- [2] R. Kembel, "The Fibre Channel Consultant: A Comprehensive Introduction," Northwest Learning Associates, Tucson, 2000.
- [3] Y. Sheffi, "The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage," MIT Press, Cambridge, 2005.
- [4] M. Bundschuh, D. Klabjan and D. L. Thurston, "Modeling Robust and Reliable Supply Chains," Working Paper, University of Illinois, Urbana-Champaign, 2003.
- [5] L. V. Snyder and Z. J. M. Shen, "Managing Disruptions to Supply Chains," *The Bridge (National Academy of Engineering)*, Vol. 36, No. 4, 2006, pp. 39-45.
- [6] S. L. Hakimi, "Optimum Location of Switching Centers and the Absolute Centers and Medians of a Graph," *Operations Research*, Vol. 12, No. 3, 1964, pp. 450-459.
- [7] S. L. Hakimi, "Optimum Distribution of Switching Centers and Some Graph Related Theoretic Properties," *Operations Research*, Vol. 13, 1965, pp. 462-475.
- [8] M. B. Teitz and P. Bart, "Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph," *Operations Research*, Vol. 16, No. 5, 1968, pp. 955-961. [doi:10.1287/opre.16.5.955](https://doi.org/10.1287/opre.16.5.955)
- [9] C. S. Revelle and R. Swain, "Central Facilities Location," *Geographical Analysis*, Vol. 2, No. 1, 1970, pp. 30-42. [doi:10.1111/j.1538-4632.1970.tb00142.x](https://doi.org/10.1111/j.1538-4632.1970.tb00142.x)
- [10] R. L. Church, "COBRA: A New Formulation for the Classic p-Median Location Problem," *Annals of Operations Research*, Vol. 122, No. 1-4, 2003, pp. 103-120. [doi:10.1023/A:1026142406234](https://doi.org/10.1023/A:1026142406234)
- [11] R. D. Carr, *et al.*, "Robust Optimization of Contaminant Sensor Placement for Community Water Systems," *Mathematical Programming*, Vol. 107, No. 1-2, 2005, pp. 337-356.
- [12] J. C. James and S. Salhi, "A Tabu Search Heuristic for the Location of Multi-Type Protection Devices on Electrical Supply Tree Networks," *Journal of Combinatorial*

Optimization, Vol. 6, No. 1, 2002, pp. 81-98.
[doi:10.1023/A:1013322309009](https://doi.org/10.1023/A:1013322309009)

- [13] R. L. Church, M. P. Scaparra and R. S. Middleton, "Identifying Critical Infrastructure, The Median and Covering

Facility Interdiction Problems," *Annals of the Association of American Geographers*, Vol. 94, No. 3, 2004, pp. 491-502. [doi:10.1111/j.1467-8306.2004.00410.x](https://doi.org/10.1111/j.1467-8306.2004.00410.x)