International Journal of Operations Research

International Journal of Operations Research, Vol. 8, No. 1, 69−75 (2011)

A Multiple-Period Facility Location Model for Large-Scale Distribution Network Design Problems with Budget and Service Level Considerations

Jenn-Rong Lin[1], Ta-Hui Yang[2,*], and Yu-Chen Huang[2]

[1] Department of Transportation Science National Taiwan Ocean University, Keelung - 20224, Taiwan

[2] Department of Logistics Management National Kaohsiung First University of Science and Technology, Kaohsiung - 81164, Taiwan

Received August 2010; *Revised December 2010*; *Accepted January 2011*

Abstract—This study aims to develop a multi-period distribution system design model that provides an integrated view of the various costs, service quality and budget concerns within the planning horizon, as well as computationally feasible methods for obtaining solutions in realistic situations. The key design decisions considered in each period are: the number and location of distribution centers in the system and the routing of shipments between distribution centers and customers. Multi-period distribution systems design requires an integrated view of facility costs and transportation costs, service quality as well as budget constraints within the planning horizon. A genetic-algorithm-based approach is proposed as the solution procedure to find the optimal sequence for locating distribution centers over the planning horizon. The quality of solutions to test problems is analyzed and compared with the optimal solutions obtained from the commercial software Lingo 11.0.
Keywords— Dynamic facility location, distribution systems design, genetic algorithms, budget constraints.

1. INTRODUCTION

The purpose of this study is to formulate and analyze a facility location decision model for multi-period distribution system design with budget and service level considerations. The key design decisions considered in each period are: the number and location of distribution centers (DC's) in the system and the routing of shipments between DC's and customers. The design decisions are made with concerns for the total logistic costs, service level (measured by coverage of retail outlets: the fraction of total demand at retail outlets that are within some specified time or distance of the nearest DC) and budgets available for locating DC's in each period. Retail outlet locations are assumed known and fixed, with varied (time-dependent) demand. The concerns in this model are long-term decisions on facility investments and the selection of transportation channels in each period.

Logistics costs (including facility costs and transportation costs) and customer responsiveness are important in designing a distribution system. One key question in designing a distribution system is locating DC's. However, there is often a trade-off between the two objectives in determining the number of DC's and their locations. A network of fewer DC's allows lower facility costs but provides lower coverage of retail demand. In addition, transportation costs may be higher because shipping distances become longer. Thus, locating DC's creates the opportunity to shift the balance among the fixed facility costs, the transportation costs and the service level.

Facility location decisions are strategic in nature. Since the facility investment required to construct facilities is usually large, facilities are expected to serve over a planning horizon. Usually, the demand, the unit transportation cost and the cost of constructing DC's may change over time. In addition, we may not have enough money to construct all the planned DC's simultaneously. Thus, it is more realistic to consider budget constraints in locating DC's and time-dependent cost parameters over the planning horizon while designing a distribution system. The decisions not only involve selecting robust DC locations to efficiently serve changing demands over time, but also when to locate DC's to achieve the required service level within budget constraints over the planning horizon.

* Corresponding author's email: yang@nkfust.edu.tw

The optimal design of a multi-period distribution system design requires an integrated view of facility costs, transportation costs, service levels and budget limits in each period. The purpose of this paper is to create a formal model that provides such an integrated view and to develop methods for obtaining solutions for the design variables in practical situations.

The remainder of this paper is organized as follows. In Section 2, we review the relevant literature. In Section 3, we present a mathematical programming formulation of the problem. The solution procedure is presented in Section 4. In Section 5, we describe an illustrative example and discuss the solution quality of the solution procedure. In Section 6, we draw some conclusions.

2. LITERATURE REVIEW

The research draws on literature in four key areas: uncapacitated facility location models (UFLP), maximal covering models, dynamic facility location models and applications of genetic algorithms to location problems. The uncapacitated facility model and maximal covering are both important in static facility location. Both models have been extensively used in practice and numerous algorithms have been proposed for their solution. Mirchandani and Francis (1990) and Daskin (1995) present surveys of application and solution procedures for both the UFLP and the maximal covering location model. The UFLP identifies the subset of potential facility locations that minimizes the cost of serving a set of demand locations. Whereas the UFLP focuses on costs, the maximal covering location problem focuses on service responsiveness. A demand location is "covered" if a facility is located within a given distance of the location (as a surrogate of service responsiveness). The maximal covering problem identifies P locations that maximize the amount of covered demand. There is often an inherent trade-off between cost and service responsiveness. That is, the lowest solution may provide poor service responsiveness and the maximal coverage solution may be expensive. Therefore, in many real world facility location problems it is important to identify a solution that represents an acceptable trade-off between these two objectives (Bhaskaran and Turnquist, 1990; Nozick, 2001).

Whereas static facility location models ignore time, dynamic facility location models incorporate the effect of the future time dimension. Since the pioneering work of Ballou (1968), researchers have been interested in dynamic facility location problems. Current et al. (1997) define two types of dynamic models: implicitly dynamic and explicitly dynamic models. In implicitly dynamic models, facilities are assumed to be opened in one period and then remain open over the planning horizon. By contrast, facilities will be opened and possibly closed over the horizon in explicitly dynamic models. Examples of implicitly dynamic models include Drezner (1995) and Hinojosa et al. (2000). Examples of explicitly dynamic models include Sweeney et al. (1976), Van Roy and Erlenkotter (1982), and Melachrinoudis et al. (2000). Mirchandani and Francis (1990), Current et al. (1997), and Owen and Daskin (1998) present reviews of dynamic facility location problems.

Genetic algorithms (GAs) were first developed by Holland (1975). Since then, GAs have been implemented in a wide variety of application areas, especially in combinatorial optimization problems. The basic idea of GAs is based on the mechanics of natural selection in biological systems. GAs use a structured but randomized way to use the information provided by existing solutions to seeking better solutions. A simple GA consists of three operators: reproduction, crossover and mutation that reflect nature's evolutionary process (Goldberg, 1989). However, applications of GAs to location problems have been relatively few. Applications of GAs to the location problems range from the p-median problem (Hostage and Goodchild, 1986, and Alp et al., 2003), to the set covering problem (Beasley and Chu, 1996), to the hub location problem (Topcuoglu et al., 2005), and to the facility location problem (Houck et al., 1996 and Jaramillo et al., 2002). A notable study that has applied GAs to solve the dynamic facility location problem is that by Ko and Evans (2007). Since there is relatively little literature published on the application of GAs to the dynamic facility location problems, there is all the more reason to apply GAs to the multi-period distribution system design problem.

3. PROBLEM DEFINITION

The problem can be summarized as follows. Given a set of potential distribution centers' locations and retail outlets with specific demand processes in each period, we would like to know where to locate distribution centers and the shipments of finished goods from each distribution center to each retail outlet in each period to ensure the specified service level within budget constraints and at the minimum total logistic cost over the planning horizon.

3.1 Model Formulation

We define the following subscripts, sets, decision variables and input parameters.

Subscripts and sets:

- $d \in D$ denotes the distribution centers (DC' s).

- $r \in R$ denotes the retailer outlets.
- $t \in T$ denotes the time periods within the planning horizon.

Decision variables:

- X_{dr}^{t} equals 1 if DC *d* is opened at the beginning of time period *t* and o otherwise.
- Z_{dr}^{t} is the volume of the product to be transported from DC *d* to retailer *r* in time period *t*.

Input parameters:

- l_{dr} is the distance from DC *d* to retailer *r*.
- ω_{dr}^{t} is the per-mile cost to ship a unit of product from DC *d* to retailer *r*.
- λ_{r}^{t} is the demand at retailer *r* in time period *t*.
- F_{d}^{t} is the fixed cost of opening DC *d* at the beginning of time period *t*.
- ψ^{t} is the budget limit allocated to opening DC's in time period *t*.
- q_{dr}^{t} equals 1 if a DC located at candidate site *d* can not cover demand at retailer *r* in the time period *t* and 0 otherwise.
- γ^{t} is the unit penalty cost for uncovered demand in time period *t*.

Based on this notation, we can develop the following mathematical formulation.

$$\text{Min} \sum_{t \in T} \left(\sum_{d \in D} \sum_{r \in R} l_{dr} \omega_{dr}^{t} Z_{dr}^{t} + \sum_{d \in D} F_{d}^{t} X_{d}^{t} + \gamma^{t} \sum_{d \in D} \sum_{r \in R} q_{dr}^{t} Z_{dr}^{t} \right) \tag{1}$$

Subject to

$$\sum_{d \in D} F_{d}^{t} X_{d}^{t} <= \psi^{t} \quad \forall d \in D, \forall t \in T \tag{2}$$

$$\sum_{d \in D} Z_{dr}^{t} = \lambda_{r}^{t} \quad \forall r \in R, \forall t \in T \tag{3}$$

$$Z_{dr}^{t} \le \sum_{t'=1}^{t} X_{d}^{t'} \lambda_{r}^{t} \quad \forall d \in D, \forall r \in R, \forall t \in T \tag{4}$$

$$Z_{dr}^{t} \ge 0 \quad \forall d \in D, \forall r \in R, \forall t \in T \tag{5}$$

$$X_{d}^{t} = \{0,1\} \quad \forall d \in D, \forall t \in T \tag{6}$$

The objective is to minimize the sum of transportation costs, facility costs and the penalty costs for uncovered demand in each period over the planning horizon. Constraints (2) ensure that facility investment in each time period can not exceed the budget available. Constraints (3) ensure that all the demand at retailers is satisfied in each time period. Constraints (4) ensure that DC's that are still closed at the beginning of each time period can not transport product to retailers. Constraints (5) ensure that the flow variables are nonnegative. Constraints (6) are an integral requirement of the location variables.

Unfortunately, an exact solution of this formulation is not computationally tractable. The key difficulties lie in the number of location variables which is the dimension of the static problem multiplied by the time periods. Since the related single-period UFLP is an NP-hard problem (Cornuejols et al., 1990) and the formulation provided above is at least as computationally complex as UFLP, we are forced to turn to heuristic solution procedures.

4. SOLUTION PROCEDURE

The solution procedure developed is a genetic algorithm heuristic for simultaneously finding the optimal system configurations over the planning horizon. The GA's solution procedure employed is simply as follows:

Generate initial population randomly
  DO WHILE stooping criteria not satisfied
    DO GA's operator WHILE new population not yet generated
      Reproduction: select the parents from the current population via roulette wheel selection for mating
      Crossover: first clone the chromosomes of the parents to the offspring and then exchange partial chromosomes
        Mutation: mutate the chromosomes of the offspring
      END DO

END DO
Report the best string as the final solution

4.1 String Encoding

Each chromosome within the GA's population represents a possible solution to the dynamic distribution systems design problem. In this problem, a chromosome consists of two strings as shown in Fig. 1. The first string represents the sequence of distribution centers to open, and the second string represents how many distribution centers are to be opened in each planning period. Putting the information provided by the two strings together, we can know how many DC's we need to open in each period. For example, the first string indicates that the sequence of distribution centers to be opened is at candidate sites 7, 2, 1, 5, 4, 3, 9, 4, 6, 10, 8 in that order. The second string means that we open two DC's in the first period and then two DC's in the second period and one in the third period, respectively. Therefore, the chromosome indicates that we locate two DC's at candidate sites 7 and 2 in the first period, two DC's at candidate sites 1 and 5 in the second period, and one DC at candidate site 4.

First String: 7 2 1 5 4 3 9 4 6 10 8

Second String: 2 2 1 0

Figure 1. A chromosome represents a solution of the problem

4.2 Calculation of Fitness Function

Because this problem is a minimization problem, we can not directly use the value of the objective function as a fitness measure of the solutions. Therefore, we need to map the objective function to a fitness function through the following transformation. The value of a fitness function for a chromosome is the value of the objective function of the worst solution in the current generation minus that of the corresponding solution.

4.3 Initial Population and the GA's Operators

The initial population is randomly generated as follows. The first string is a random permutation of all potential DC sites. The second string consists of a series of numbers representing the number of DC's to open in each period over the planning horizon. The number of DC's to open in each period is randomly picked up from zero to the maximal number of DC's allowed to open within the budget limit in that period. The two procedures are repeated until the specified population size is met.

The initial population is randomly generated. The GA's operators are used to select parent strings and to generate the next generation. The parents are selected via roulette wheel selection which allows the fitter individual strings to have a higher chance of being selected as a parent string to the mating pool. The idea behind this method is that each current string in the population has a roulette wheel slot size in proportion to its fitness. Selecting a string of a population to be a parent can be viewed as a spin of the wheel with the winning string being the one in whose slot the spinner stops. It is defined as follows:

1. Sum up the fitness over all individuals in the current population
2. Calculate the percentage of the population's total fitness for each individual
3. Select an individual as a parent based on the percentage of the population's total fitness

Since each chromosome is encoded as a binary string, we simply use a single point crossover and simple mutation that flips a single bit with a low probability. The algorithm terminates when it reaches the maximum number of generations.

5. ILLUSTRATIVE EXAMPLE

We demonstrate our overall algorithm on a small example excerpted from Daskin (1995). There are fifty retailer outlet locations and each is considered to be a potential DC location. The fifty retailer outlets are some state capitals plus some other large cities. The locations of the retail outlets and the average retail demand over three planning periods are shown in Figure 2 where the size of circle represents the volume of demand. The delivery cost from DC's to retailers is assumed to be \$0.6 per unit-mile. The construction cost of a DC is estimated to be \$10 million and the budget available for constructing DC's is \$30 million in each period. Finally we assume that the coverage distance (for service quality requirements) is 400 miles and the uncovered penalty cost is \$1,600 per unit.

We implement the algorithm in MATLAB on a Pentium IV 1200 MHz PC with the following parameters: population size = 50, crossover rate =1.0, mutation rate =0.1 and the number of generations =1000. Figure 3 shows the optimal sequence of locating DC's. In the first period, we locate three DC's in Los Angeles, CA; St. Louis, MO; and Trenton, NJ. We then locate three DC's in Jacksonville, FL; Houston, DA; and Portland, OR in the second period and three DC's in Chicago, IL; Phoenix, AZ; and Wichita, KS, respectively.

In order to investigate the solution quality of the solution procedure, four additional test problems are created which are similar to the illustrative example, but with different demand patterns, DC construction costs and uncovered demand penalty costs. For the illustrative example and the four additional test problems, it is possible to solve the problems optimally with the optimization solver LINGO 11.0. Table 1 compares the solution from the procedure with the optimal solution solved by LINGO 11.0. Notice that for all test problems the solution procedure identifies the optimal solution at least 1 time out of ten trials and within 4% of the optimum on average. Computationally, the solution procedure is quite good and holds substantial promise for the solution of the large problem.

Figure 2. Location of sites for example problem

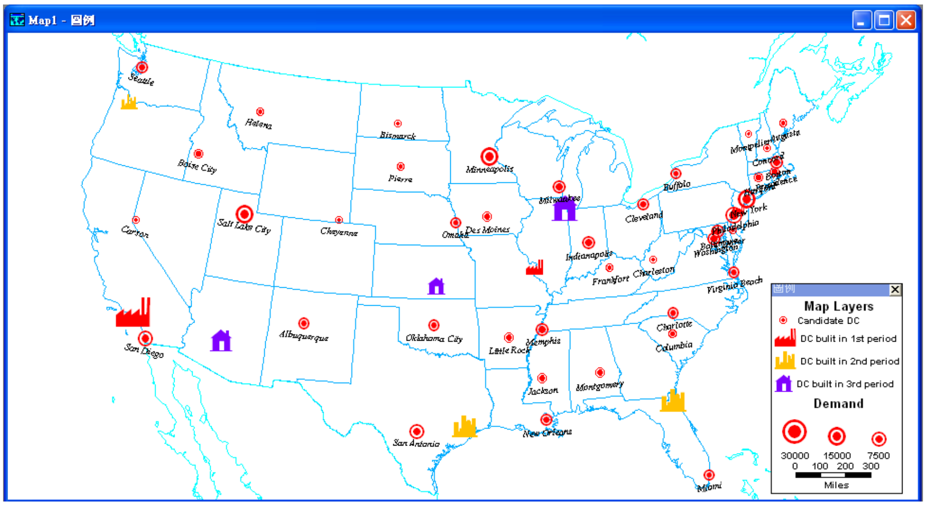

Figure 3. Location of open DC's in each period

Table 1. Solution Quality

Test Problem	**Optimal Solution**	**Solution Procedure**			
		Average	Best Found	Worst Found	Time Optimal Found
Example	304.0	314.7	304.0	329.5	4
Problem 1	158.5	161.0	158.5	167.4	5
Problem 2	540.9	562.4	540.9	593.0	1
Problem 3	209.8	210.5	209.8	212.0	7
Problem 4	494.3	507.1	494.3	520.3	2

6. CONCLUSIONS

Multi-period distribution system design requires an integrated view of facility costs, transportation costs, service quality as well as budget constraints within the planning horizon. This paper develops a model that provides an integrated view of various costs, service quality and budget concerns, as well as a computationally feasible solution method for obtaining a solution in realistic situations. The solution procedure developed is a genetic algorithm approach for finding the optimal sequence of system configurations. In order to evaluate the quality of solution produced by the procedure, a series of test problems is solved. For all the test problems, the solution procedure identifies the optimal solution at least 1 time out of ten trials and identifies solutions within a 4% optimum on average. Computationally, the procedure appears to produce very good solutions, and it holds significant promise for the solution to large problems.

Further enhancements would be useful in at least the following directions. The current structure of the system under study only consists of two-echelons: distribution centers and retail outlets. A more general structure of system should include at least one more echelon: production plants. Second, few dynamic facility location models have found practical applications. It might therefore be useful to undertake a real-world case study to demonstrate the usefulness of the multi-period distribution design model.

7. ACKNOWLEDGMENTS

This work was partially undertaken with support from the National Science Council, Taiwan, under contract NSE 92-2211-E-327-003. This support is greatly acknowledged.

REFERENCES

1. Alp, O., Erkut, E. and Drezner, Z. (2003). An Efficient Genetic Algorithm for the P-Median Problem, *Annals of Operations Research*, 122: 21-42.
2. Ballou, R.H. (1968). Dynamic Warehouse Location Analysis, *Journal of Marketing Research*, 5: 271-276.
3. Beasley, J.E. and Chu, P.E. (1996). A Genetic Algorithm for the Set Covering Problem, *European Journal of Operational Research*, 94: 392-404.
4. Bhaskaran, S. and Turnquist, M.A. (1990). Multiobjective Transportation Considerations in Multiple Facility Location, *Transportation Research Part A*, 24: 139-148.
5. Current, J., Ratick, S., and Revelle, C. (1997). Dynamic Facility Location When the Total Number of Facilities is Uncertain: A Decision Analysis Approach, *European Journal of Operational Research*, 110: 567-609.
6. Daskin, M. S. (1995). *Network and Discrete Location: Models, Algorithms and Applications*, John Wiley and Sons, New York.
7. Drezner, Z. (1995). Dynamic Facility Location: The Progressive P-Median Problem, *Location Science*, 3(1): 1-7.
8. Goldberg, D. E. (1989). *Genetic Algorithms in Search*, Optimization & Machine Learning, Addison Wesley, New York.
9. Hinojosa, Y., Puerto, J., and Fernandez, F.R. (2000). A Multiperiod Two-echelon Multicommodity Capacitated Plant Location Problem, *European Journal of Operational Research*, 123: 271-291.
10. Holland, J.H. (1975). *Adaptation in Neural and Artificial Systems*. Univ. Michigan Press, Ann Arbor, MI.
11. Hosage, C.M. and Goodchild, M.F. (1986). Discrete Space Location-Allocation Solutions for Genetic Algorithms, *Annals of Operations Research*, 6: 35-46.
12. Houck, C.R., Joines, J.A., and Kay, M.G. (1996). Comparison of Genetic Algorithms, Random Restart and Two-Opt Switching for Solving Large Location-Allocation Problems, *Computers & Operations Research*, 23: 587-596.
13. Jaramillo, J.H., Bhadury, J., and Batta, R. (2002). On the Use of Genetic Algorithms to Solve Location Problems, *Computers & Operations Research*, 29: 761-779.
14. Ko, H.J. and Evans, G.W. (2007) A Genetic Algorithm-based Heuristic for the Dynamic Integrated forward/Reverse Logistics network for 3PLs, *Computers & Operations Research*, 34: 346-366.
15. Melachrinoudis, E. and Hokey Min (2000). The Dynamic Relocation and Phase-out of a Hybrid, Two-echelon Plant/Warehousing Facility: A Multiple Objective Approach, *European Journal of Operational Research*, 123: 1-15.
16. Mirchandani, P. B. and Francis, R. X. (1990). *Discrete Location Theory*, John Wiley and Sons, New York.
17. Nozick, L.K. (2001). The Fixed Charge Facility Location Problem with Coverage Restrictions, *Transportation Research Part E*, 37: 281-296.
18. Owen, S. H. and Daskin, M.S. (1998). Strategic Facility Location: A Review, *European Journal of Operational Research*, 111: 423-447.

19. Sweeney, D.J. and Tatham, R.L. (1976). An Improved Long-run Model for Multiple Warehouse Location, *Management Science*, 22(7): 748-758.
20. Topcuoglu, H., Corut, F., Ermis, M., and Yilmaz, G. (2005). Solving the Uncapacitated Hub Location Problem Using Genetic Algorithms, *Computers & Operations Research*, 32: 967-984.
21. Van Roy, T.J. and Erlenkotter, D. (1982). A Dual-Based Procedure for Dynamic Facility Location, *Management Science*, 28(10): 1091-1105.