A Multi-objective Evolutionary Approach for Cloud Service Provider Selection Problems with Dynamic Demands

Hsin-Kai Chen, Cheng-Yuan Lin, and Jian-Hung Chen(✉)

Dept of Computer Science and Information Engineering,
Chung-Hua University, Hsin-Chu, Taiwan
jh.chen@ieee.org

Abstract. This paper describes a multi-objective evolutionary approach for solving cloud computing service provider selection problems with dynamic demands. In this investigated problem, not only the service purchase costs and transmission costs of service providers are different, but the demands of service requests also change over the given periods. The objective of this problem is to select a number of cloud service provider while optimizing the total service distance, the total number of serviced demand points, the total service purchase costs, and total transmission costs simultaneously in the given continuous time periods. A multi-objective genetic approach with a seeding mechanism is proposed to solve the investigated problems. Four trail benchmark problems are designed and solved using the proposed multi-objective evolutionary algorithm. The results indicate that the proposed approach is capable of obtaining a number of non-dominated solutions for decision makers.

Keywords: Cloud computing · Multi-objective optimization · Dynamic optimization · Evolutionary algorithms

1 Introduction

With the rapid development of computing hardware, high-speed network, web programming, distributed and parallel computing, and other storage technologies, cloud computing has recently emerged as an effective reuse paradigm, where hardware computing power, software functionality, and other computing resources are delivered as integrated services through Internet [1]. There are many global and local commercial cloud service providers, offering various kinds of delivered services such as Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Recently, the advantages and features of cloud services has arisen the interests of digital entertainment/media/content suppliers to integrate cloud computing services into their content delivery networks [2].

Consider a national-wide area with a number of service request points, the requests at each point usually changes in time; and within this area, a number of cloud service providers with different locations and pricing options of services are available for

A.I. Esparcia-Alcázar et al. (Eds.): EvoApplications 2014, LNCS 8602, pp. 841–852, 2014.
DOI: 10.1007/978-3-662-45523-4_68

chosen. From the point view of digital entertainment/media content suppliers, it is an important issue to select suitable cloud computing service providers, which can deliver their contents to massive customers rapidly and smoothly. Therefore, maximizing some expected Quality-of-Service (QoS) indictors and minimizing services related costs are crucial considerations for decision makers. As a result, considering the requirements of content supplier and the conditions of cloud service providers, we formulated such problems to multi-objective dynamic p-median problems in this paper.

The classical p-median problem consists of selecting p facilities in a given space which minimizes the total costs of serving m demand points at a time. P-median problem is prominent combinatorial optimization NP-hard problem in location science and cluster analysis [3-9]. Many exact and heuristic approaches have been proposed for solving p-median problems [3][8][9]. In traditional approaches, the planning of service facility centers usually considers the demand of consumers as constant values. However, it is not true in the real world applications, because the demands of consumers may change by environments and time. The dynamic p-median problem is applicable to all situations modeled by the standard p-median problem whenever demand changes over time in a predictable way.

In this paper, a multi-objective p-median model with dynamic demands which optimizes the total QoS distance, the total number of serviced demand points, the total service purchase costs, and the total network transmission costs is investigated. Considering four different geographical features, we propose an efficient approach based on genetic algorithms for content providers to determine the selection of service providers in different periods and satisfying the dynamic demands of customers. The proposed approach can also provide decision-makers a set of non-dominated solutions for the selection processes.

This paper is organized as follows: Section 2 describes the investigated dynamic p-median problem and multi-objective optimization. Section 3 describes the mathematical model of the investigated problem. Section 4 presents the proposed multi-objective genetic algorithm MOGA for solving investigated problems. Section 5 gives the experimental results and analysis of the proposed algorithm. Section 6 concludes our paper.

2 Related Work

2.1 P-median Problems

The classical p-median problem consists of locating p facilities (medians) in a given space (e.g. Euclidean space) which minimizes the total costs of serving m demand points, where the pair-wise cost of servicing each point from all facilities is given. Each demand point is only served by a single facility and services to demand points are not combinable [3-10].

Exact methods for solving p-median problems include linear programming approaches, dual-based algorithms. However, these exact methods suffer from the curse of dimensionality since the computation costs of calculating all demand points' expectations over all possible future combinations increases exponentially in the number of demand points. Many heuristic approaches have been proposed to solve p-median problems, including greedy heuristic, variable neighbor decomposition search, cooperative parallel variable neighborhood search, and Lagrangian-surrogate heuristic. Modern meta-heuristics have been applied to solve p-median problems as well [8], such as tabu search approaches, simulated annealing approaches and genetic algorithms approaches.

Recently, considering the real-world conditions, various models of p-median problems are proposed in the literature, including stochastic p-median problems, progressive p-median problems [3], dynamic p-median problems, and bi-objective p-median problems [9].

2.2 Multi-objective Evolutionary Optimization

Assume the multi-objective functions are to be minimized. Mathematically, MOOPs can be represented as the following vector mathematical programming problems

$$\mathit{Minimize} \quad F(Y) = \{F_1(Y), F_2(Y), ..., F_i(Y)\}, \tag{1}$$

where Y denotes a solution and $f_i(Y)$ is generally a nonlinear objective function. Pareto dominance relationship and some related terminologies are introduced below. When the following inequalities hold between two solutions Y_1 and Y_2, Y_2 is a non-dominated solution and is said to dominate Y_1 ($Y_2 \succ Y_1$):

$$\forall i : F_i(Y_1) > F_i(Y_2) \wedge \exists j : F_j(Y_1) > F_j(Y_2). \tag{2}$$

When the following inequality hold between two solutions Y_1 and Y_2, Y_2 is said to weakly dominate Y_1 ($Y_2 \succeq Y_1$):

$$\forall i : F_i(Y_1) \geq F_i(Y_2). \tag{3}$$

A feasible solution Y^* is said to be a Pareto-optimal solution if and only if there does not exist a feasible solution Y where Y dominates Y^*, and the corresponding vector of Pareto-optimal solutions is called Pareto-optimal front.

By making use of Pareto dominance relationship, multi-objective evolutionary algorithms (MOEAs) are capable of performing the fitness assignment of multiple objectives without using relative preferences of multiple objectives. Thus, all the objective functions can be optimized simultaneously. As a result, MOEA seems to be an alternative approach to solving the investigated service provider selection problems on the assumption that no prior preference and domain knowledge is available [10-11].

3 Cloud Service Selection Problems with Dynamic Demands

In this paper, the investigated dynamic service provider selection problem (DSPSP) is to select p service providers from n service providers in each season, in order to satisfy the dynamic demands of m service requests from end-users. The following conditions are assumed in this problem:

1) Each service provider has different pricing options for purchasing services and network transmission.
2) Although contents can be deliver to anywhere though internet, end-users still expects no delays during network transmission. Therefore, each service provider has a pre-assumed maximum QoS distance.
3) The number of demand points that a service provider can service is unlimited.
4) The Euclidean distance is used to calculate the distances between demand points and points of service provider.
5) Each demand point can only serviced by a nearest point of service provider within the maximum QoS distance.
6) In order to satisfying the dynamic demands, content supplier may select p different service providers in the next following season.

The investigated problem can be formulated to multi-objective p-median problems with dynamic demands. The objectives of DSPSP are while optimizing four competing objective functions: the total QoS distance, the total number of serviced demand points, the total service purchase costs, and the total network transmission costs.

3.1 Problem Notations

i , j : $i \in \{1,2,3,\ldots m\}$, $j \in \{1,2,3,\ldots,n\}$.
m : The total number of demand points.
n : The total number of service provider points for selection.
L_i : The index of demand points, $L_i = i$.
S_j : The index of the service provider points. Service providers points usually colocate with some demand points, therefore $S_j \in \{L_1, L_2, \ldots L_m\}$.
D_j : The maximum QoS distance of the service provider point j.
T : The total service periods.
t_j : The time period that the service provider S_j served, $0=t_1<t_2<\ldots<t_p<t_{p+1}<T$.
d_{ij} : The distance between L_i and S_j.
md_{ij} : The nearest distance of the demand point L_i between the nearest service provider point, $md_{ij}=min\{d_{ij}\}$.
$w_i(t)$: The demanding function of the demand points L_i at time t, $0 \leq t<T$.
w_{ij} : The total demanding amount of the demand point L_i from time t_j to time t_{j+1},

$$w_{ij} = \int_{t_j}^{t_{j+1}} w_i(t)\, dt.$$

A_j : The network transmission cost of the service provider point S_j per demand unit.
C_j : The monthly service purchase cost of the service provider point S_j.
X_i : The serviced index of the demand point L_i. If the demand point service L_i is serviced within the maximum QoS distance of a provider point, then $X_i = 1$, otherwise $X_i = 0$.

Z_j : The selection index of the service provider point S_j. If the service provider point S_j is chosen and serves demand points in the specific time period, then $Z_j = 1$, otherwise $Z_j = 0$.

3.2 Problem Objectives

1. Minimization of QoS distance
2. In the classical p-median problem, the demands in each demands points are usually considered to a constant. However, considering the real-world applications, demands are known to be changed dynamically. Given the demanding function of each demand points, the QoS distance of each demand to its nearest service provider points can be expressed as follows:

$$Minimize\, F_1 = \sum_{j=1}^{n}\sum_{i=1}^{m} w_{ij} \times md_{ij} \times X_i \times Z_j. \tag{4}$$

3. Minimization of network transmission cost

Considering the cloud computing environments, the costs of network transmissions between service provider points and demand points are not fixed. Given the network transmission cost of each service point per time unit, the transmission costs of each facility can be expressed as follows:

$$Minimize\ F_2 = \sum_{j=1}^{n}\sum_{i=1}^{m} w_{ij} \times A_j \times X_i \times Z_j. \tag{5}$$

4. Minimization of service purchase cost

In additional to the network transmission cost, the service purchase cost on a specific service provider point is also an important factor for content suppliers, because the service cost in different service provider point are different. Given the service purchase cost for each service provider points, the total service purchase costs of selected service provider points can be expressed as follows:

$$Minimize\ F_3 = \sum_{j=1}^{n} C_j \times Z_j. \tag{6}$$

5. Maximum of total number of serviced demand points

Because different service providers has different QoS distance, therefore the number of demand points that a service provider points may serviced could be different. Given the maximum QoS distance of each service provider, the number of serviced demand points can be calculated as follows:

$$Maximize\ F_4 = \sum_{i=1}^{m} X_i. \tag{7}$$

3.3 An Illustrative Example

An example is given here to explain our mathematical formation. Assumed that a content supplier plans to select three service provider points (p=3) from six providers (n=6) within twelve months (T=12), in order to service ten demanding points (m=10). The maximum QoS *Dj* is 3 for all the service provider points. The coordination, demanding function of demand points, the service purchase costs and transmission costs of service provider points are listed in Table 1. Assumed a selection plan for four seasons is determined (as shown in Table 2), three service provider S_2, S_3, S_6 are select in the first season, and finally three service provider S_1, S_3, S_5 are select in the fourth season.

Take the selection plan of Season 4 for example, the total amount of each demand points during Season 4 can be calculated, as shown in Table 3. The distance of each demand point to different service provider points can be calculated, as shown in Table 4. The demand points with Dj are marked as bold. Hereafter, according to all the tables, the objective functions in Season 4 can be calculated, F_1 = 10.12242, F_2 = 1507.5, F_3 = 1650, F_4 = 8.

Table 1. The information of demand and service points L_i ,S_j

L_i	S_j	coord.	$w_i(t)$	A_j	C_j
L_1	S_1	(1,8)	10+6t	1	500
L_2	S_2	(2,5)	3+4t	1	700
L_3		(0,9)	16+2t	1	
L_4		(10,2)	25+3t	1	
L_5	S_3	(4,5)	50-2t	1	700
L_6	S_4	(3,7)	99-3t	1	450
L_7	S_5	(12,3)	6+7t	1	450
L_8		(6,16)	24+4t	1	
L_9		(2,10)	10+10t	1	
L_{10}	S_6	(8,4)	5+5t	1	500

Table 2. representation of four selection plan for four seasons

Season 1	Season 2	Season 3	Season 4
2, 6, 3	3, 6, 4	5, 4, 3	3, 1, 5

Table 3. The total amount of demands in season 4, according to the selection plan

	t = 0~3	t = 3~6	t = 6~9	t = 9~12
L_1	57	111	165	219
L_2	27	63	99	135
L_3	57	75	93	111
L_4	88.5	115.5	142.5	169.5
L_5	141	123	105	87
L_6	283.5	256.5	229.5	202.5
L_7	49.5	112.5	175.5	238.5
L_8	90	126	162	198
L_9	75	165	255	345
L_{10}	37.5	82.5	127.5	172.5

Table 4. The distance of each demand point to selected service provider points in quarther 4

	$S_3(=L_5)$	$S_1(=L_1)$	$S_5(=L_7)$
L_1	4.24264	**0**	12.083
L_2	**2**	3.16228	10.198
L_3	5.65685	**1.41421**	13.4164
L_4	6.7082	10.8167	**2.23607**
L_5	**0**	4.24264	8.24621
L_6	**2.23607**	**2.23607**	9.84886
L_7	8.24621	12.083	**0**
L_8	11.1803	9.43398	14.3178
L_9	5.38516	**2.23607**	12.2066
L_{10}	4.12311	8.06226	4.12311

4 The Proposed Multi-objective Genetic Algorithm

In this section, the proposed multi-objective genetic algorithm to find a selection plan within four seasons for DSPSP is described.

4.1 Chromosome Representation

A chromosome has gene information for solving the problem in DSPSP. In the proposed approach, each chromosome of has p genes. When a season is finished, the non-dominated solutions will be selected as seed chromosomes for the initial population of the next season. The chromosome can be regarded as a selection plan for a season.

4.2 Fitness Assignment

We use a generalized Pareto-based scale-independent fitness function (GPSIFF) considering the quantitative fitness values in Pareto space for both dominated and non-dominated individuals [10]. GPSIFF makes the best use of Pareto dominance relationship to evaluate individuals using a single measure of performance. The used GPSIFF is briefly described below. Let the fitness value of an individual Y be a tournament-like score obtained from all participant individuals by the following function:

$$F(X) = Np - Nq + c., \tag{8}$$

where Np is the number of individuals which can be dominated by the individual Y, and Nq is the number of individuals which can dominate the individual Y in the objective space. Generally, a constant c can be optionally added in the fitness function to make fitness values positive. c is usually set to the number of all participant individuals.

4.3 Procedure of MOGA

The procedure of MOGA is written as follows:

Input: population size Npop, recombination probability pc, mutation probability pm, the number of maximum generations Gmax. Current Season Index q=1.
Output: The optimum solutions ever found in P.
Step 1: *Initialization* Randomly generate chromosomes to fill in the population P until Npop individuals are reached. Each chromosome is consists of p genes for a season.
Step 2: *Evaluation* For each individual in the population, compute all objective function values F_1, F_2, F_3.and F_4.
Step 3: *Fitness Assignment* Assign each individual a fitness value by using the equation (8) GPSIFF.
Step 4: *Selection* Select Npop individuals from the population to form a new population using the binary tournament selection without replacement,.
Step 5: *Recombination* Perform the uniform crossover operation with a recombination probability p_c.
Step 6: *One Point Mutation* Apply the one point mutation operators to each gene with a mutation probability p_m. If the mutated gene is duplicated with other genes in the same chromosome, mutate the gene again.
Step 7: *Termination* test If the maximum generations have reached, store all the non-dominated solutions in season q, and then go to Step 8. Otherwise, go to Step 2.
Step 8: *Seeding* q=q+1. If q>4, stop the algorithm. Otherwise, select and copy non-dominated solutions to the initial population of the next season. If the number of non-dominated solutions is greater than the population size Npop, randomly delete solutions until the population size is equal to Npop. Then, go to Step 1.

5 Result and Discussions

5.1 Simulation Environment and Parameter Settings

In this paper, four benchmarks are designed for experiments, as shown in Fig. 1. Each problem has different distribution of demand points on different grid sizes, described as follows:

- Circle: 100 demand points and 36 service providers on a 18*18 grid. The number of providers to be chosen p=10, and the maximum QoS distance D_j=2.2.
- Rectangle: Square with empty space. 100 demand points and 36 service providers on a 16*16 grid. The number of providers to be chosen p=10, and the maximum QoS distance D_j=3.

- Square: 100 demand points and 36 service providers on a 110*110 grid. The number of providers to be chosen p=10, and the maximum QoS distance D_j=10.
- Triangle: 100 demand points and 36 service providers on a 14*14 grid. The number of providers to be chosen p=10, and the maximum QoS distance D_j=2.

Ten service providers will be select for each season. The total number of season is 4. The parameter settings of MOGA are listed as follows: population size N_{pop}=100, recombination probability p_c=0.9, mutation probability p_m=0.1, the number of maximum generations G_{max}=100. Fifteen independent runs are conducted for each problem.

5.2 Discussions

For each benchmarks, 30 independent runs are conducted using MOGA with seeding mechanism and MOGA without mechanism. Figure 2-5 use boxplot to depict the values F_1 of non-dominated solutions in solving the circle benchmark at different seasons. From these figures, it shows that seeding mechanism can help MOGA obtains better solutions and converge faster. Figure 5-8 use boxplot to depict the values of F_1, F_2, F_3, and F_4 of non-dominated solutions in solving the circle benchmark at the Season 4. Figure 9-12 use boxplot to depict the values of F_1, F_2, F_3, and F_4 of non-dominated solutions in solving the Rectangle benchmark at the Season 4. Due to the page limit, the results of Square and Triangle are not shown in this paper. The results indicate that the proposed MOGA is capable of solving DSPSP and optimize four objectives simultaneously, considering different geographic distribution of demand points.

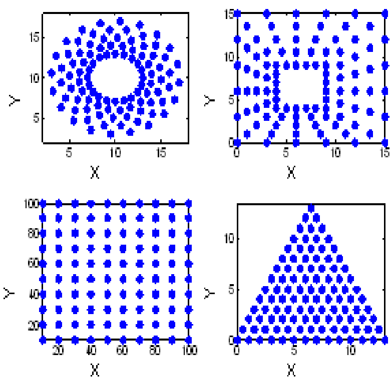

Fig. 1. Distributions of demand points in four benchmark problems

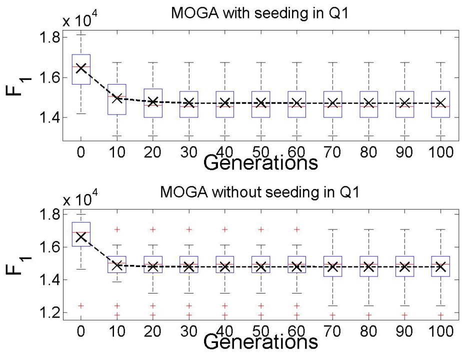

Fig. 2. F_1 of non-dominated solutions for circle benchmark in Season 1

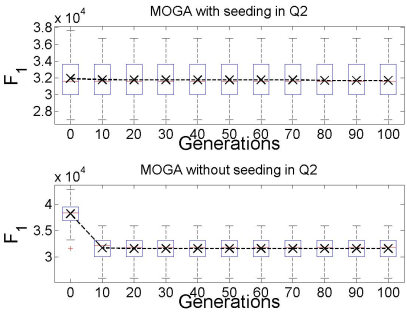

Fig. 3. F_1 for circle benchmark in Season 2

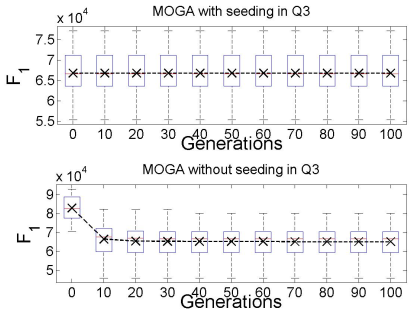

Fig. 4. F_1 for circle benchmark in Season 3

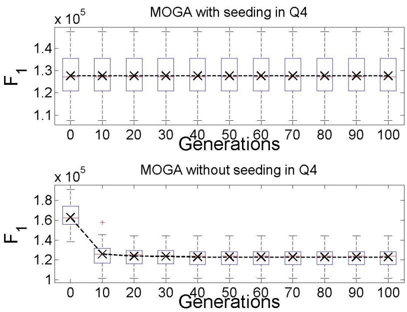

Fig. 5. F_1 for circle benchmark in Season 4

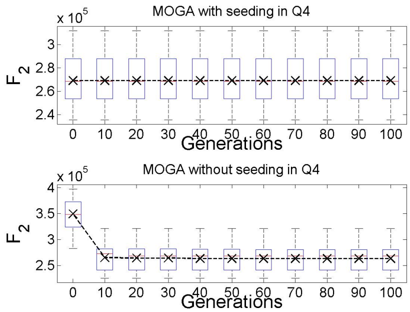

Fig. 6. F_2 for circle benchmark in Season 4

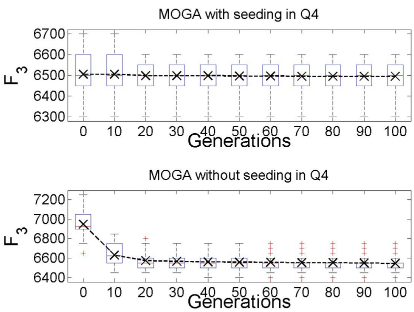

Fig. 7. F_3 for circle benchmark in Season 4

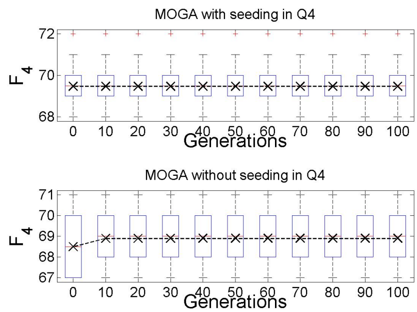

Fig. 8. F_4 for Rectangle benchmark in Season 4

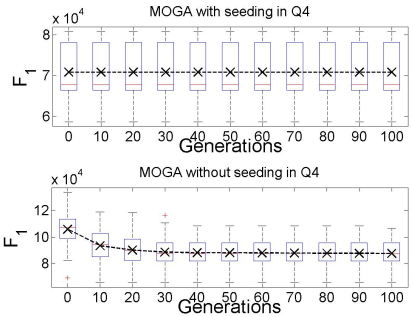

Fig. 9. F_1 for Rectangle benchmark in Season 4

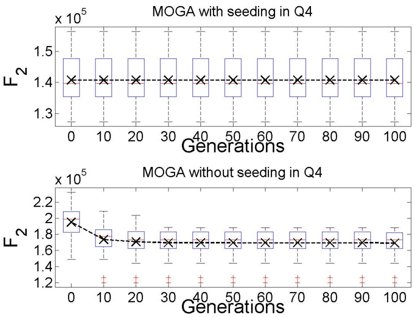

Fig. 10. F_2 for Rectangle benchmark in Season 4

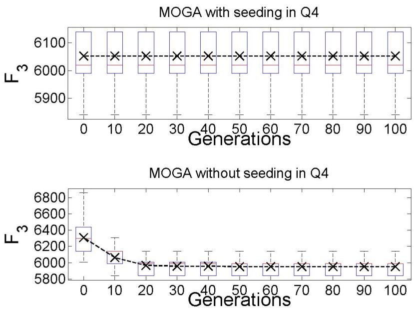

Fig. 11. F_3 for Rectangle benchmark in Season 4

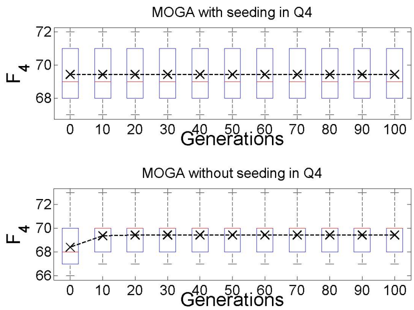

Fig. 12. F_4 for Rectangle benchmark in Season 4

6 Conclusions

In this paper, a multi-objective evolutionary approach is proposed to solve dynamic service provider selection problems. Experimental results demonstrated the proposed approach is capable of optimizing the QoS distance, the total network transmission cost, the total service purchase cost, and the total number of demands points simultaneously. Moreover, the proposed approach can provide mission planers a set of non-dominated solutions for construction plan of service facilities. Our future work is to apply our approach in solving some real cases.

References

1. Foster, I., Zhao, Y., Raicu, I., Lu, S.: Cloud Computing and Grid Computing 360-Degree Compared. In: Proceeding of Grid Computing Environments Workshop, GCE 2008, pp. 1–10, November 12–16 (2008)
2. Li, Y., Shen, Y., Liu, Y.: Utilizing Content Delivery Network in Cloud Computing. In: Proceeding of 2012 International Conference on Computational Problem-Solving (ICCP), pp. 137–143 (October 2012)

3. Drezner, Z.: Dynamic Facility Location: The Progressive *p*-median Problem. Location Science **3**(1), 1–7 (1995)
4. Own, S.H., Daskin, M.S.: Strategic Facility Location: A Review. European Journal of Operational Research **111**, 423–447 (1998)
5. Wesolowsky, G.O.: Dynamic Facility Location. Management Science **19**(11), 1241–1248 (1973)
6. Wesolowsky, G.O., Truscott, W.G.: The Multiperiod Location-Allocation Problem with Relocation of Facilities. Management Science **22**(1), 57–65 (1975)
7. Francisco, S.D.G., Maria, E.C.: A Heuristic Approach for the Discrete Dynamic Location Problem. Location Science **6**, 211–223 (1998)
8. Pullan, W.: A population based hybrid metaheuristic for the p-median problem. In: Proceedings of IEEE Congress on Evolutionary Computation, pp. 75–82 (June 2008)
9. Arroyo, J.E.C., dos Santos Soares, M., dos Santos, P.M.: A GRASP heuristic with Path-Relinking for a bi-objective p-median problem. In: Proceedings of 10th International Conference on Hybrid Intelligent Systems (HIS), pp. 97–102 (August 2010)
10. Ho, S.-Y., Shu, L.-S., Chen, J.-H.: Intelligent Evolutionary Algorithms for Large Parameter Optimization Problems. IEEE Transaction on Evolutionary Computation **8**(6), 522–541 (2004)
11. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strengthen Pareto approach. IEEE Transaction on Evolutionary Computation **3**(4), 257–271 (1999)