The Huff Versus the Pareto-Huff Customer Choice Rules in a Discrete Competitive Location Model

Pascual Fernández[1(✉)], Blas Pelegrín[1], Algirdas Lančinskas[2], and Julius Žilinskas[2]

[1] Department Statistics and Operations Research, University of Murcia, Murcia, Spain
{pfdez,pelegrin}@um.es

[2] Institute of Mathematics and Informatics, Vilnius University, Vilnius, Lithuania
{algirdas.lancinskas,julius.zilinskas}@mii.vu.lt

Abstract. An entering firm wants to compete for market share in a given area by opening some new facilities selected among a finite set of potential locations. Customers are spatially separated and other firms are already operating in that area. In this paper, we analyse the effect of two different customers' behavior over the optimal solutions, the Huff and the Pareto-Huff customer choice rules. In the first, the customer splits its demand among all competing facilities according to its attractions. In the second, the customer splits its demand among the facilities that are Pareto optimal with respect to the attraction (to be maximized) and the distance (to be minimized), proportionally to their attractions. So, a competitive facility location problem on discrete space is considered in which an entering firm wants to locate a fixed number of new facilities for market share maximization when both Huff and Pareto-Huff customer behavior are used. In order to solve these two models, a heuristic procedure is proposed to obtain the best solutions, and it is compared with a classical genetic algorithm for a set of real geographical coordinates and population data of municipalities in Spain.

1 Introduction

Probably, the most important decision for a firm that competes with other firms to provide goods or services to the customers in a given geographical area, is where to locate its facilities. Depending on location space, facility attraction, customer behavior, demand function, decision variables, etc., different location models and solution procedures have been proposed (see for instance survey papers [6,11]). The entering firm is focused in the determination of the optimal locations for the new facilities in order to maximize its market share or profit, taking into account the customers' behavior and that have to compete with other pre-existing facilities belonging to different firms for customers demand. Usually, it is assumed that the customers choose the nearest facility to be served, but

O. Gervasi et al. (Eds.): ICCSA 2018, LNCS 10961, pp. 583–592, 2018.
https://doi.org/10.1007/978-3-319-95165-2_41

on real problems, customers take into account some other characteristics of the facilities, in addition to the distance.

Since the attraction model proposed by Huff [9], different customer choice rules have been used to estimate the market share captured by the competing facilities. The attraction of a facility is defined as the quotient between facility quality (which depends on its characteristics) and a non-negative non-descending function of the distance between the customer and the facility. The most common customer choice rules are the Huff (or proportional or probabilistic) and the binary (or deterministic) (see [12]). In the Huff case, customers patronize all the facilities in proportion to facility attraction (see for instance [3,13]), and in the binary case, each customer patronizes the most attractive facility (see [1,7]).

When Huff customer choice rule is considered, customers will patronize very distant facilities even when more attractive facilities are much closer, although the captured demand by these facilities is small. To avoid these allocations and adjust the model more to reality, a modified Huff rule is considered, the so called Pareto-Huff customer choice rule [10], which has hardly been used in the literature. In this case, a customer will patronize a more distant facility only if it is more attractive, then distant facilities will be selected by customers only if no facility exists that is both closer and at least so attractive. So, on Pareto-Huff model, each customer splits its demand among the facilities that are Pareto optimal with respect to the attraction (to be maximized) and the distance (to be minimized), proportionally with their attractions.

In this paper we will consider these two customer choice rules, the Huff and the Pareto-Huff, will analyse its influence on the optimal solutions, and will propose a heuristic procedure to be solved them since these two models are not linear programming problems. To solve both problems, one of the algorithms proposed by Fernández *et al.* (2017) has been used, the one based on ranking and distance, because it can be used to solve any discrete competitive location problem, and since it do not require to have analytical expression of the objective function, but only be able to evaluate objective value at any solution candidate. To check performance of the proposed heuristic algorithm, it is necessary to know the optimal solution of the problems in order to compare it with the solution given by the heuristic algorithm. For small size data, a complete enumeration algorithm to obtain optimal solutions has been used for both models. For bigger size data, a classical genetic algorithm is used to obtain good solutions to be compared with solutions provided by the here proposed heuristic algorithm.

The reminder of the paper is organized as follows: Sect. 2 consists of description of the competitive location problems and its formulations, Sect. 3 is devoted to presentation of the heuristic algorithm, and Sect. 4 includes the description and discussion of the experimental investigation of the proposed algorithm; finally, conclusions are presented in Sect. 5.

2 Discrete Location Models

An entering firm wants to open new facilities in an area where similar facilities of other competing firms are already present. For simplicity, it is considered that all pre-existing facilities belong to the same firm, the competitor. Customers are aggregated to geographic demand points in order to make the problem computationally tractable (see [5] for demand aggregation), and its demand is fixed and known.

The following general notation is used:
Indices:

i, I index and set of demand points (customers)
j, k indeces of facilities

Data:

w_i demand at i
q_j quality of facility j
d_{ij} distance between demand point i and facility j
a_{ij} attraction that demand point i feels for facility j
L set of candidate locations for the new facilities
C set of pre-existing facilities of competetitor/s
r number of new facilities to be located

Variables:

X set of locations for the new facilities

2.1 Model with the Huff Customer Choice Rule

When Huff customer choice rule is considered, each customer splits his demand over all facilities in the market proportionally with his attraction to each facility (additive attraction). Note that in this model it is not necessary to know if the pre-existing facilities owns or not to different firms, since the demand of each customer i is split between all facilities, regardless of the firm to which they belong.

Usually, the attraction that a customer i feels for a facility j is defined as $a_{ij} = \frac{q_j}{g(d_{ij})}$, where $g(d_{ij})$ is a non-decreasing non-negative and convex function of the distance between customers and facilities. In this paper, it has been considered that $g(d_{ij}) = 1 + d_{ij}$ for simplicity.

If $M_H(X)$ denote the market share captured by the entering firm when Huff customer choice rule is used and its new facilities are located at X, the problem can be formulated as:

$$Max\{M_H(X) = \sum_{i\in I} w_i \frac{\sum_{j\in X} a_{ij}}{\sum_{j\in X} a_{ij} + \sum_{k\in C} a_{ik}} : |X| = r, X \subset L\}. \qquad (1)$$

that is a nonlinear optimization problem.

2.2 Model with the Pareto-Huff Customer Choice Rule

In this new model, the demand of each customer i will be split between all Pareto optimal facilities PH_i with respect to quality and distance, proportionally with their attraction. For any customer i, facilities belonging to PH_i are non-dominated facilities with respect to quality and distance, i.e., for any $j \in PH_i$ there doesn't exist any other facility with at least quality q_j and closer to i than j (see Fig. 1 for an example).

If $M_{PH}(X)$ denote the market share captured by the entering firm when its new facilities are located at X, and the Pareto-Huff customer choice rule is considered, the problem can be formulated as:

$$Max\{M_{PH}(X) = \sum_{i \in I} w_i \frac{\sum_{j \in PH_i \cap X} a_{ij}}{\sum_{j \in PH_i} a_{ij}} : |X| = r, X \subset L\}. \tag{2}$$

which, as the Huff model, it is a nonlinear optimization problem.

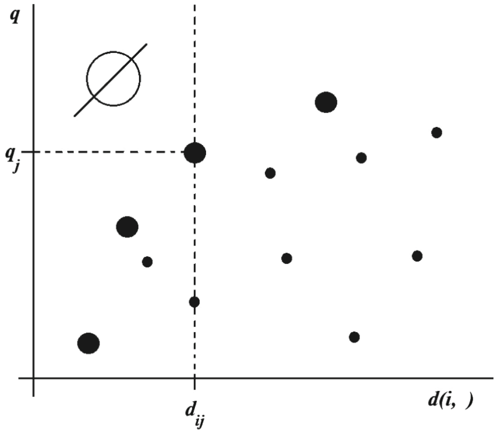

Fig. 1. Example of Pareto optimal facilities PH_i for customer i (big points).

3 Ranking-Based Discrete Optimization Algorithm

A heuristic algorithm has been developed to solve the location problems described in the previous section, which do not guarantee determination of the optimal solution, but rather its approximation. The algorithm is based on a single agent random search in the neighborhood of the best solution found so far, and its only requirement is to have availability to evaluate objective value at any possible solution. The Discrete Optimization Algorithm based on Ranking and Distance (RD) (see [4]) begins with a subset of location candidates X as an initial variable, which is considered as the best known (but not necessarily optimal) solution of the problem being solved. A new solution X' is generated by in turn taking location from X and changing them to another one with probability $\pi_c = 1/r$, thus ensuring that a single facility will be changed in average. The new facility l, in case of change, is selected from $L' = L \setminus (X \cup X')$ as

$$x_i' = \begin{cases} l \in L \setminus (X \cup X') & \text{if } \xi_i < 1/r \\ x_i & \text{otherwise} \end{cases} \tag{3}$$

where ξ_i is a random number uniformly generated over the interval $[0, 1]$, and $i = 1, 2, \ldots, r$.

Each candidate location $l_i \in L$, has different probability to be selected. This probability is based on the rank r_i of the candidate location l_i, and a geographical distance $d(l_i, x_z')$ between the candidate location l_i and the location x_z'; here x_z' is a location already forming solution X' and is expected to be changed $(z = 1, 2, \ldots, r)$, and can be expressed as

$$\pi_i^{rd} = \frac{r_i}{d(l_i, x_z') \sum_{j=1}^{|L|} \frac{r_j}{d(l_j, x_z')}} \tag{4}$$

At the beginning of the algorithm, all candidate solutions have unit ranks $r_i = 1$, forall $i = 1, 2, \ldots, |L|$. Later ranks are automatically adjusted according to the success and failures in improvement of the best known solution by selecting a particular candidate location. If the newly generated solution X' improves the best solution found so far X, i.e. $M(X') > M(X)$, where $M(\cdot)$ stands for the objective function, then

(1) the ranks of all candidate locations in X' are increased by one, and
(2) the ranks of all candidate locations that form X, but do not form X' are reduced by one.

Otherwise, if $M(X') \leq M(X)$, then the ranks of all candidate locations which form unsuccessfully generated solution X', but do not form the best known solution X, are reduced by one. If any of the ranks reaches zero, then ranks are increased by one.

If the new solution X' improves X, then X is changed by X'. Such a process is continued till a stopping criterion, which is usually based on the number of function evaluations, is satisfied.

4 Experimental Investigation

The proposed algorithm RD has been experimentally investigated by solving both considered models using real geographical data of coordinates and population of 6960 municipalities in Spain, numbered in decreasing order with respect to its population, which will be considered as demand points and its demand equal to the population, with total demand around 38,5 millions of inhabitants. The distances between demand points and facilities have been calculated in kilometers using Haversine distance [14], and the attraction that demand point i feels for facility j has been taken as $a_{ij} = \frac{q_j}{1+d_{ij}}$.

For all experiments, it has been considered that the number of pre-existing facilities is equal to 30 (the most populated, nodes from 1 to 30), and its quality values have been randomly generated in the interval $[30, 70]$. All new facilities for the entering firm is supposed to have the same quality equal to 35, 45, 55 and 65. Some combinations of parameters $(r, q, |L|)$ are considered when $r = 5, 10$, $q = 35, 45, 55, 65$, and $|L| = 500, 1000$.

The goodness of the proposed algorithm has been evaluated by the quality of the best solution X found. The best solution found by the proposed algorithm has been approximated using 10000 function evaluations for each instance. Due to its stochastic nature, the algorithm has been run for 100 independent runs and average results have been considered. The results obtained by RD have been compared with the results obtained by a Genetic Algorithm (GA) [2,8] with the population of 100 individuals, uniform crossover with the rate of 0.8 and mutation rate of $1/r$; GA has been run for 100 generations, thus performing 10000 function evaluations in total. All experiments have been run in a PC with Pentium IV Processor, 3.2 GHz and 3 GB RAM.

For the smaller data set, parameters $(5, q, 500)$, a Complete Enumeration Algorithm (CEA) has been used to obtain optimal solutions for both models. In this case, the quality of the best solution is the ratio between its objective value and the objective value of the optimal solutions provided by CEA, and to measure its quality, the probability to achieve the optimal solution with different error has been evaluated.

For the Huff model, the results show that RD provides better probability to achieve optimal solutions, independent of problem instance. The proposed heuristic algorithm guarantees to find the optimal solution with 1% error, independent of problem instance, whilst GA can provide a guarantee optimal solution with error of 4% for any instance (see Table 1).

Table 1. Huff model: results for parameters $(5, q, 500)$.

r	q	$\mid L \mid$	0%	1%	2%	3%	4%	5%
RD								
5	35	500	0,79	1,00	1,00	1,00	1,00	1,00
5	45	500	0,84	1,00	1,00	1,00	1,00	1,00
5	55	500	0,82	1,00	1,00	1,00	1,00	1,00
5	65	500	0,83	1,00	1,00	1,00	1,00	1,00
GA								
5	35	500	0,00	0,91	0,93	1,00	1,00	1,00
5	45	500	0,38	0,86	0,92	0,96	1,00	1,00
5	55	500	0,00	0,83	0,94	0,99	1,00	1,00
5	65	500	0,30	0,83	0,93	0,97	1,00	1,00

For Pareto-Huff model the results are similar. RD provides better probability to achieve optimal solution for all instances, guarantees to find the optimal solution with 2% error, while GA cannot guaranty optimal solution with error less than 5% for any instance (see Table 2).

Table 2. Pareto-Huff model: results for parameters $(5, q, 500)$.

r	q	$\mid L \mid$	0%	1%	2%	3%	4%	5%
RD								
5	35	500	0,76	1,00	1,00	1,00	1,00	1,00
5	45	500	0,98	0,98	1,00	1,00	1,00	1,00
5	55	500	0,77	1,00	1,00	1,00	1,00	1,00
5	65	500	0,95	1,00	1,00	1,00	1,00	1,00
GA								
5	35	500	0,00	0,17	0,68	0,99	1,00	1,00
5	45	500	0,03	0,28	0,45	0,85	0,97	0,97
5	55	500	0,10	0,80	0,90	0,93	0,96	0,99
5	65	500	0,03	0,16	0,64	0,73	0,91	0,98

For parameters $(5, q, 500)$, $(10, q, 500)$ and $(10, q, 1000)$, where $q = 35, 45, 55, 65$, solutions provide by RD and GA algorithms are compared (see Table 3). For the Huff model, the new facilities for all the parameters are located, in general, in points with greater demand, however, for the Pareto-Huff model, for q = 34, 45, the new facilities are not located in the points of greater demand, but in points with lower demand near the previous ones, and for q = 55, 65, the results are similar to the Huff model, although the locations for the new facilities are not the same. Regarding the total demand captured by the new facilities in both models, in the Huff model, greater market share is obtained for q = 35, 45, while for q = 55, 65, with the Pareto-Huff model, greater market share is captured.

The quality of the solution given by each algorithm has been measured by the ratio of average of utility function of the 100 runs, and the best known value of utility function obtained by any of the algorithms. Each experiment has also been ran for 10000 function evaluations and repeated 100 times. Results show that RD solves each instance better than GA independent on the parameters and the customers choice rule. Furthermore GA was more sensitive to the parameter of the problem. The quality of solutions obtained by RD for any set of parameters is always between 0,990 and 1, and by GA between 0,944 and 0,996 (see Table 4).

Table 3. Best solutions found by using RD algorithm.

	$(r, \|L\|)$	q	X	$M(X)$ (millions)
Huff	(5, 500)	35	1, 2, 3, 4, 6	5,38
		45	1, 2, 3, 4, 6	6,48
		55	1, 2, 3, 4, 6	7,46
		65	2, 3, 4, 5, 6	8,35
	(10, 500)	35	1, 2, 3, 4, 5, 6, 8, 14, 32, 35	8,36
		45	1, 2, 3, 4, 5, 6, 8, 14, 29, 32	9,95
		55	1, 2, 3, 4, 5, 6, 8, 14, 29, 32	11,3
		65	1, 2, 3, 4, 5, 6, 7, 8, 14, 29	12,6
	(10, 1000)	35	1, 2, 3, 4, 5, 6, 8, 14, 32, 35	8,36
		45	1, 2, 3, 4, 5, 6, 8, 14, 29, 32	9,95
		55	1, 2, 3, 4, 5, 6, 8, 14, 29, 32	11,3
		65	1, 2, 3, 4, 5, 6, 7, 8, 14, 29	12,6
Pareto-Huff	(5, 500)	35	21, 40, 55, 108, 112	2,55
		45	6, 14, 15, 55, 108	3,65
		55	1, 3, 4, 6, 12	10,2
		65	1, 3, 4, 12, 22	11,2
	(10, 500)	35	12, 21, 40, 53, 55, 101, 108, 112, 149, 377	4,26
		45	6, 12, 14, 15, 21, 55, 101, 108, 112, 149	5,78
		55	1, 3, 4, 5, 6, 8, 12, 14, 15, 23	14,6
		65	1, 3, 4, 5, 6, 7, 8, 12, 14, 22	16,6
	(10, 1000)	35	12, 21, 40, 53, 55, 101, 108, 112, 149, 377	4,26
		45	6, 12, 14, 15, 21, 55, 108, 112, 149, 621	5,78
		55	1, 3, 4, 5, 6, 8, 12, 14, 15, 23	14,6
		65	1, 3, 4, 5, 6, 7, 8, 12, 14, 22	16,6

Table 4. Results for parameters $(5, q, 500)$, $(10, q, 500)$ and $(10, q, 1000)$.

	$(r, \mid L \mid)$	$q = 35$		$q = 45$		$q = 55$		$q = 65$	
		RD	GA	RD	GA	RD	GA	RD	GA
Huff	(5,500)	0,999	0,996	0,999	0,995	0,999	0,995	1,000	0,994
	(10,500)	0,999	0,989	0,999	0,987	1,000	0,980	1,000	0,986
	(10,1000)	0,998	0,981	0,996	0,982	0,997	0,973	0,998	0,978
Pareto-Huff	(5,500)	0,999	0,984	1,000	0,981	0,999	0,991	1,000	0,980
	(10,500)	0,998	0,975	0,998	0,971	0,994	0,959	0,998	0,956
	(10,1000)	0,995	0,969	0,996	0,959	0,990	0,963	0,992	0,944

5 Conclusions

In this paper two discrete competitive facility location models for an entering firm have been considered depending if Huff or Pareto-Huff customers choice rule is considered. A heuristic algorithm has been proposed to solve these non-linear location problems, based on the ranking and the distance between candidate locations and customers [4]. To check the performance of the proposed algorithms we have used real geographical coordinates and population data of 6960 municipalities in Spain, and have compared the solutions generated with the best feasible solution found, and in the particular, for small size data, a complete enumeration algorithm has been used to obtain the optimal solution. Computational experiments prove that the proposed algorithm is always better than GA, and RD obtains always the best solution for all parameters of the problem. The quality of the solution obtained with RD is in average 0.998 when all the results for Huff and Pareto-Huff models are considered.

In conclusion, this heuristic algorithm, RD, can be a good option to solve discrete competitive facility location problems when an entering firm wants to locate new facilities in the market, and customer choice rule is Huff or Pareto-Huff.

Acknowledgments. This research has been supported by the Ministry of Economy and Competitiveness of Spain under the research project MTM2015-70260-P, and the Fundación Séneca (The Agency of Science and Technology of the Region of Murcia) under the research project 19241/PI/14, and also by a grant (No. MIP-051/2014) from the Research Council of Lithuania.

References

1. Campos, C.M., Santos-Peñate, D.R., Moreno, J.A.: An exact procedure and LP formulations for the leader-follower location problem. Top **18**(1), 97–121 (2010)
2. Davis, L.: Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York (1991)
3. Drezner, T., Drezner, Z.: Finding the optimal solution to the Huff based competitive location model. Comput. Manag. Sci. **1**(2), 193–208 (2004)
4. Fernández, P., Pelegrín, B., Lančinskas, A., Žilinskas, J.: New heuristic algorithms for discrete competitive location problems with binary and partially binary customer behavior. Comput. Oper. Res. **79**, 12–18 (2017)
5. Francis, R.L., Lowe, T.J., Tamir, A.: Demand point aggregation for location models. In: Hamacher, H., Drezner, Z. (eds.) Facility Location: Application and Theory, pp. 207–232 (2002)
6. Friesz, T.: Competitive networks facility location models: a survey. Pap. Reg. Sci. **65**, 47–57 (1998)
7. Hakimi, L.: Location with spatial interactions: competitive locations and games. In: Drezner, Z. (ed.) Facility Location: A Survey of Applications and Methods, pp. 367–386 (1995)
8. Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975)

9. Huff, D.: Defining and estimating a trade area. J. Mark. **28**, 34–38 (1964)
10. Peeters, P.H., Plastria, F.: Discretization results for the Huff and Pareto-Huff competitive location models on networks. TOP **6**, 247–260 (1998)
11. ReVelle, C., Eiselt, H., Daskin, M.: A bibliography for some fundamental problem categories in discrete location science. Eur. J. Oper. Res. **184**(3), 817–848 (2008)
12. Serra, D., Eiselt, H.A., Laporte, G., ReVelle, C.S.: Market capture models under various customer-choice rules. Environ. Plan. B: Plan. Des. **26**, 741–750 (1999)
13. Serra, D., Colomé, R.: Consumer choice and optimal location models: formulations and heuristics. Pap. Reg. Sci. **80**(4), 439–464 (2001)
14. Sinnott, R.W.: Virtues of the haversine. Sky Telescope **68**(2), 159 (1984)