Computational Research Progress in Applied Science & Engineering

ISSN 2423-4591

CRPASE Vol. 05(04), 114-117, December 2019

A Method to Reach New Goals from Past Experiences: Cased Based Reasoning Systems in Transportation

Rohollah Khaani Ghorotkhor[a*], Reza Ganjbakhsh Sanati[b]

[a] *Department of Mathematics, Islamic Azad University, Zabol Branch, Zabol, Iran*
[b] *Department of Mathematics, University of Guilan, Rasht, Iran*

Keywords

Shortest path,
CBR method,
Dijkstra and Floyd-Warshal method.

Abstract

In this paper, applying the Case-Based Reasoning (CBR) method which is based on previously experienced problems to solve new similar ones, we study a special kind of well-known problem hawker vendor. We provide a database of special information extracted from an online taxi services company about daily trips and drivers. Then, using a function defined by a new recorded trip's data and considering the shortest path, we try to guess the driver who accepts the trip.

1. Introduction

The cost of supplying raw materials and component parts through vendors forms a significant part of the goods's overall expense.

Actually, the cost of purchasing raw materials and outside services occupies, in average, 70% of the value of factories's final product [9]. This percent can even reach to 80 in high-tech companies [4].

Therefore, it is very important for companies to select suitable vendors as such as selecting the right vendors is affective in reducing costs and increasing the competitiveness of companies, choosing inappropriate vendors can cause degradation of financial and operational position of the companies.

The purchase managers, in the companies in which supplying raw materials, buying process and selecting the proper vendors are repetitous as a daily activity, use previously buying as an experience to decide about the same new buying and choosing the appropriate vendors for each new goods.

The case-based reasoning (CBR), using soloutions ways of preriovsely problems to solve new similar problems, can be a proper tool to make appropreiate decision.

In this paper, we study the issue of selecting the proper vendors by extending a model based on CBR method in approximetely fuzzy environment.

In the following, we review studies on some research topics such as the related concepts of fuzzy theory, using for provided model one of the most known methods to solve the vendor selecting problems is the linear weighting method. In this method, each criterion of vendors selecting is weight assigned. Indeed, vendors are evaluated by these criterias and specified by a score in this model. Finally, the weighted score of each vendor is considered to select the appropreiate seller and to prioritization [6, 11,16].

Also, mathematical programming methods have been studies by researchers for vendors selecting problems. One of the first researchs in these topics, 1974, discusses about mixed integer programming (MIP) to apply the maximum discount suggested by seller [14]. The methods based on linear programming recognize the proper sellers and allocate them on order by minimizing purchase costs, considering some limitations such as vendors capacity, buyer's demand and service quality of the sellers [1,13]. Moreover, linear and nonlinear ideal programming have been applied in the study of selecting the proper vendors. In these studies, price targets, service level, delivery and quality are considered as the goal functions [3, 15]. Some other methods such as "Data Envelopment Analysis (DEA)" and "Analytical Hierarchy process (AHP)" have been investigated to measure the seller's performance [7, 19]. Also, some combination methods have been used to solve the problem [2, 10, 18].

2. Background

Here, a short review of the background material that our work depends opun is presented.

Dijkstra, 1959, and Floyd-Warshal, 1996, are the most wellknown methods for shortest path finding algorithms.

Consider a net $N(V,A)$ with the "nodes" set $V=\{1,2,3,\dots,n\}$, $(|V|=n)$ and "edjes" set $\mathrm{A}=\{(\mathrm{i},\mathrm{k}) : \mathrm{i},\mathrm{k}\in \mathrm{V},\ \mathrm{i}\neq \mathrm{k}\}$. The condition $\mathrm{i}\neq\mathrm{k}$, indeed, expresses that the

* Corresponding Author:
E-mail address: Rohollah.Khaani@gmail.com – Tel, (+98) 9038315239

Received: 15 October 2019; Accepted: 10 December 2019

graph in the problem has no rings. Dijkstra algorithm is the best one in all methods to find shortest path. In a more general view, when the graph can have some rings, Floyd-Warshal algorithm plays the role. This algorithm is based on a 4-steps process in which two matrixes D_j and R_j $(j = 0,1,2,\dots,n)$ will be calculated. D_j saves the lowest costs values in itself and R_j, the shortest path values between the origin and the destination. Actually, the minimal cost and the shortest path between any two nodes in a net with rings are resulted by n+1 steps. It must be notioned in this algorithm that although the calculations are simple and easy, the high number and the long time of the calculations is an significant concepts. In what follows, the Floyd-Warshal algorithm is explained.

Step 1.

Let n be the number of nodes in the net or the graph. Define two n×n matrixes D_j and R_j $(j = 0,1,2,\dots,n)$. Note that every "j" is named a "step" itself.

Step 2.

For j=0, the elements d_{ik} of the matrix D_0 are:

$$\boldsymbol{d_{ik}} = \begin{cases} d_{ik} & \text{if there is a direct path from the node i to node k} \\ 0 & \text{if i = k} \\ \infty & \text{if there is no direct path from the node i to node k.} \end{cases}$$

And the elements r_{ik} of the matrix R_0 are:

$$\boldsymbol{r_{ik}} = \begin{cases} k & \text{if there is a direct path from the node i to the node k} \\ \dots & \text{if i = k} \\ \dots & \text{if there is no direct path from the node i to the node k.} \end{cases}$$

Step 3.

For j=1,2,3,…,n, these elements are:

$$d_{ik} = \begin{cases} d_{ik} & \text{if i = k = j} \\ min\{d_{ik}\ ,\ d_{ij} + d_{jk}\} & else. \end{cases}$$

And

$$r_{ik} = \begin{cases} k & \text{if } i = k = j \\ k & \text{if } d_{ik} \le d_{ij} + d_{jk} \\ j & \text{if } d_{ik} > d_{ij} + d_{jk}. \end{cases}$$

Step 4.

Continue the step3 untill the matrixes D_n and R_n be completely obtined.

3. The Case-Based Reasoning (CBR)

The method CBR, indeed, follows the way people behave dealing with new issues in which uses the experiences of solving past problems as a guide to solve the new ones. Solving problems based on CBR method acts in 4 steps:

1) Recover the same case as the new one in new problems.
2) Using the same recovered problem to give suggested answer for new problems.
3) Review of suggested answer in the case of confliction between the condition of new problems and te recovered one.
4) Keeping the new obtained results for future uses.

Every case consists of two parts. The first part specifies relevant problems and the second one, expresses the answer of the raised problem. The specification of a case determines the properties of it's components and assigned values of these featurs shows the situation of the case.

4. Problem Definition

Prediction of accepting an online trip ordered by a driver, according to a predefined database, is the general frame work of our research in this paper.

The affective factors for accepting a trip by adriver which we considered are:

1) The proposed of the trip (in kilometers).
2) The distance of the trip (in kilometers).
3) Traffic jam amount in the route.
4) Type of the trip divided into two outing city and intra urban trips.
5) Stopping rate in the route.
6) Type of passengers divided into two family and single modes.

According these factors, every trip is assigned by an array containing the trip's information and the accepting driver's code. Measuring the similarity of the new problem with the previously solved ones is the most important part in the implementation of a CBR problem. Actually, offering the similarity indicators or a matching fits and using these indexes, we look for the nearest neighbor for the new problem. By neighbors, we mean the members of the database which we created. Specially, in the online travel order problem, these members are, indeed, the arrays containing information of the trip. Estimating the rate of matching (bit to bit) by the Eculidean meters is one of the most common methods to measure the similarity in the arrays. Therefore, the bits of every trip's array which reports qualitative concepts most be numerical coded.

According trips reception's factors by drivers, the factors such as type of the trip, type of passengers and amount of traffic jam are qualitative factors. So, in the case of dual mode factors such as trip's types (outing city or intra urban) and passenger's types (family or single) we choose codes 0,1. Also, reporting the amount of trffic jam we use the numbers:

0, for no traffic jam.
1, for light traffic jam.
2, for semi-heavy traffic jam.
3, for heavy traffic jam.

5. Adaptive Modeling to Retrieve Similar Trips

After making a database of registered travel's characters, we have to offer a model to adapt the new ordered trip with previouse ones in the database. According to these concepts, we show by N, the new ordered trip and by Ω, the created database. Also, $P \in \Omega$ means that P is a registered trip. Now, we introduce the following notions:

r_{Ni} expresses the content of the number i bit of the new trip N.

r_{Pi} expresses the content of the number i bit of the ordered trip P.
w_i shows the weight of the number i factor in the calculations.

Now define the function $sim(N, P)$ as

$$sim(R, I) = \begin{cases} \left(\sqrt{\sum_{i=1}^{6} w_i (r_{Ni} - r_{Pi})^2}\right)^{-1} & \text{if } r_{Ni} \neq r_{Pi}\,, \quad \text{for some i} \\ 1 & \text{if } r_{Ni} \neq r_{Pi}\,, \quad \text{for some i.} \end{cases}$$

Computing the impact coefficient of each factor wi , we have gathered the information on 20 trips in Table 1. Then, applying "unsupervised clustering method", We run K-means algorithm for two modes K=3 and K=5.

We used evaluation algorithm "j48" to evaluate which mode works best. The results are outlined in Table 2 and Table 3 as follows.

Table 2. Applying algorithm "j48" for K=3

TP Rate	0.75
FP Rate	0.137
Precision	0.729
Recall	0.75
E-measure	0.737
MCC	0.64
ROC Area	0.857
PRC Area	0.791

Table 3. Applying algorithm "j48" for K=5

TP Rate	0.85
FP Rate	0.137
Precision	0.733
Recall	0.85
E-measure	0.784
MCC	0.753
ROC Area	0.948
PRC Area	0.845

It is obvious that the accuracy of K-means algorithm for K=5 is greater than the othere mode K=3, by comparing the two tables. Hence, the results of clustering with K=5 were selected to appoint the supervisor fild. Now the algorithm “feature selection” is applied to measure the importance of each feature. Table 4 shows the results as follows:

Table 4. Impact factor of features

Feature	Impact Pachir
Price	0.996
Kilometer	1
Traffic	0.475
t-type	1
Stop	0.994
P-type	1

Price: the proposed tariff for the trip
Kilometer: the dotance of the trip
Traffic: traffic jam amount in the route
t-type: type of the trip, auting city or intra urban trips
Stop: Stopping role in the route
P-type: Type of passengers, family or single modes.

Table 1. Data base

Code of The Reciepting driver	Traffic jam amount of the route	Stopping rate in the trip (minutes)	Kind of the passenger	Kind of the trip	The distance of the trip (kilometer)	Price of the trip
4261	1	0	0	1	7	1.5
4233	2	0	1	1	5	2.2
4125	0	0	1	0	25	4
4250	1	0	0	1	9	3
4168	1	15	1	1	10	5.5
4143	1	0	0	0	23	5
4283	0	5	0	1	6	1.5
4262	1	0	1	1	7.5	2
4142	3	0	1	1	6	4.5
4125	1	0	0	0	75	6.2
4283	0	0	1	1	8	2
4253	3	10	0	1	12	5
4262	1	0	1	1	9.5	3
4125	0	30	0	0	100	8
4261	0	0	1	1	8	2
4196	1	20	0	1	17	5.2
4257	0	0	0	0	32	4.5
4283	1	0	1	1	10	3
4172	0	5	1	0	41	5.5
4291	0	0	0	1	7	2

Now, the pseudo-code for the adopting algorithm as follows:

/* The matrix A is our data base and the array W is the array of the featur's impact factors */

Float array C

Float array B

/* User enters array B */

For $(\mathrm{i} = 1;\ \mathrm{i} \le 20;\ \mathrm{i}++)$

{

$C[i] = SIM(A_i\ , B[i])$

/* A_i is the i-th row of the matrix A */

}

$max = C[1]$

For $(k = 2; k \le 20;\ k++)$

{

$if\ \ (C[1]\ < C[k])$

$max = C[k]$

$index = k$

}

Print $A[index][7]$.

/* $A[index][7]$ is the ID-number of the driver pridected to accept the trip B */ .

6. Conclusions

Choosing the right option is one of the most important issues for support managers of companies. There is always the risk of making the wrong decision on the issues that we are faced with the right option prediction. Because these decisions are entirely based on past experiences, designing and implementing a proper CBR system and carefully choosing the similarity function based on the defined attributes are essential.

References

[1] T.F. Anthony, F.P. Buffa, Strategic purchase scheduling, Journal of Purchasing and Materials Management 13 (1977) 27–31.

[2] Bargegol, M. Nikookar, R.V. Nezafat, E.J. lashkami, A.M. Roshandeh, Timing Optimization of Signalized Intersections Using Shockwave Theory by Genetic Algorithm, Computational Researc Progress in Aplied Science & Engineering 1 (2015) 160–167.

[3] F. P. Buffa, W. M. Jackson, A goal programming model for purchase planning, Journal of Purchasing and Materials Management 3 (1983) 27–34.

[4] T.T. Burton, JIT/Repetitive sourcing strategies: tying the knot with your suppliers, Production and Inventory Management Journal 4 (1988) 38–41.

[5] S. W. Changchien, M. C. Lin, Design and implementation of a case-based reasoning system for marketing plans, Expert Systems with Application 28 (2005) 43–53.

[6] S. D. Cooper, A total system for measuring of performance, Journal of Purchasing and Materials Management 1 (1977) 22–26.

[7] L. Easton, J.D. Murphy, J.N. Pearson, Purchasing performance evaluation: with data envelopment analysis, European Journal of Purchasing & Supply Management, 8 (2002) 123–134.

A. A. Gaballa, Minimum cost allocation of tenders, Operational Research Quarterly 3 (1974) 398.

B. Ghobadian, A. Stainer, T. Kiss, A computerized vendor rating system, Proceedings of the First International Symposium on Logistics, The University of Nottingham, Nottingham, UK, (1993) 321–328.

[8] S.H. Ghodsypour, C. O'Brien, A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming, International Journal of Production Economics 56-57 (1998) 199–212.

[9] R.E. Mazurak, S.R. Rao, D.W. Scotton, Spreadsheet software application in purchasing, Journal of Purchasing and Materials Management (1985) 8–16.

[10] S. Montani, R. Bellazzi, L. Portinale, G. d'Annunzio, S. Fiocchi, M. Stefanelli, Diabetic patients management exploiting case-based reasoning techniques, Computer Methods and Program in Biomedicine 3 (2000) 205–218.

A. C. Pan, Allocation of order quantity among suppliers, Journal of Purchasing and Materials Management 3 (1989) 36–39.

[11] G. Schmidt, Case-based reasoning for production scheduling, International Journal of Production Economics 56-57 (1998) 537–546.

[12] D. Sharma, W. C. Benton, R. Srivastava, Competitive strategy and purchasing decision, Proceedings of the 1989 Annual Conference of the Decision Sciences Institute, (1989) 1088-1090.

[13] E. Timmerman, An approach to vendor performance evaluation, Journal of Purchasing and Materials Management, winter (1986) 2–8.

A. Varma, N. Roddy, ICARUS: Design and deployment of a case-based reasoning system for locomotive diagnostics, Engineering Application of Artificial Intelligence 6 (1999) 681–690.

[14] M. Vosoogh, A. Addeh, An Intelligent Power Prediction Method for Wind Energy Generation Based on Optimized Fuzzy System, Computational Researc Progress in Aplied Science & Engineering 5 (2019) 34–43.

A. A. Weber, A data envelopment analysis approach to measuring vendor performance, Supply Chain Management 1 (1996) 28–39.