IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

Solving Prioritized Multi-Objective TSP Using Genetic Algorithm

Ammar Al-Dallal

Computer Engineering Dept., Ahlia University, Manama, Kingdom of Bahrain

Abstract: Genetic algorithm has been successfully adopted to solve combinatorial problems. One of which is the Travelling Salesman Problem (TSP). One of the applications of TSP is when there is a trade off between delivering goods to customers using shortest path so that it is beneficial for the service provider, and delivering it based on customer's priority so it is beneficial for the service receiver. In this paper, a multi-objective TSP is proposed to balance between shortest path and high priority using genetic algorithm. This work is featured by proposing a new fitness function to evaluate different solutions during the process of selection and crossover. The experiment is conducted by altering the factors associated with both path length and priority. The results show that better solution is achieved when more weight is assigned to the priority than when assigned to the path length.

Keywords: Genetic algorithm, Travelling Salesman Problem, mutli-objective TSP, crossover, fitness function.

I. INTRODUCTION

Genetic Algorithm (GA) is one of the evolution based algorithms which became an important approach to solve complex problems. GA is known to be one of the best approaches adopted to solve optimization problem when little knowledge about the problem that needs to be optimized is required. It is characterized by the highly parallel, random and self-adaptive algorithm which has many merits over traditional methods such as global optimization. It uses selection, crossover and mutation operators to solve optimization problems by a survival of the fittest idea.

Genetic algorithms (GAs) have been successfully applied to solve many combinatorial problems, including several types of the Traveling Salesman Problem (TSP).

TSP is one of the typical NP completeness problems in combinational optimization. It is simply described as follows. A salesman has to visit a number of cities only once, and then return home. The distances between any pair of cities are given at first and it is required to find the shortest path. The problem is to find in which order should the cities be visited to minimize the distance travelled? There are two kinds of TSP, one is symmetric TSP whe.n the distance travelled from city v_i to city v_j is equal to the distance travelled from city v_j to city v_i, and another is asymmetric TSP when the distance travelled from city v_i to city v_j is not equal to the distance travelled from city v_j to city v_i. In the paper, symmetric TSP is considered. Suppose that there are set of cities numbered 1,2,.., n, and the distance between two stochastic cities i and j is $d[i,j](i \neq j)$, the problem is to find a shortest close route passing through all the cities only once.

In math, the problem may be stated as follows [Yuan]. Given n cities, if V denotes the set of cities as $V = \{v_1, v_2, v_3, \ldots, v_n\}$, and T denotes a permutation as $T = (t_1, t_2, t_3, \ldots, t_n\}$, where $t \in V$ (i=1,2,3, …,n}, and denoted $t_{n+1} = t_1$, then TSP is given

$$\min \quad L = \sum_{i=1}^{n} d_{t_i,t_{i+1}} \qquad (1)$$

For symmetric TSP, $d_{i,j} = d_{j,i}\ \ \forall i, j = 1, 2, 3, \ldots n$.

The rest of this paper is organized as follows: Section two lists a related work of TSP that is implemented using GA. Section three describes the proposed technique to implement the proposes prioritized TSP. Section four describes the conducted experiment. The results are analysed in Section five. Finally, Section six concludes the work and highlight points of improvements.

II. RELATED WORK

Many researchers and experts have worked to solve different kinds of problems which are similar to TSP. Examples of such problems are listed here.

(Josep and Auwatanamongkol, 2000) applied multi-objective genetic algorithm to perceiving the actual content of an image during image parsing. The parsing method is used to identify multiple instances of objects of different classes and sizes. The classification of instances can be done accurately by using feature representation. In order to find a

subset of image segments that would best form an instance for a class of objects, a multi-objective genetic algorithm is applied. It is applied to maximize support vector machines score and size of the instances. Since genetic algorithm can identify a single instance for each class of objects, the authors proposed a crowding genetic algorithm for multiple optimal solutions. (Jazayeri and Sayama, 2015) proposed polynomial-time deterministic algorithm to produce a solution for the traveling salesperson problem. The cites in this algorithm are ranked according to their priorities which are calculated using a power function of means and standard deviations of their distances from other cities. The process if performed by connecting cities with higher priorities to other cities with higher priorities as neighbors. This algorithm is applied in a 20 sets of TSBLIB ranges in size between 51 and 493. The results are compared with the best-known solution and it achieved error of max 3.73% from best solution.
One variant of TSP is solving tour planning problem. (Abbaspour and Samadzadegan, 2011) adapted two genetic algorithms (GA) to solve tour planning problem in which the routes between two points of interests are part of a multimodal transportation network. One GA is developed for tour planning and the other is developed to find the multimodal shortest path. The proposed GA used variable length chromosomes. The results are applied on a data set of cities consists of 400 different requests.
Another application of TSP is the Multi-Objective TSP. (Shi Lianshuan & Li Zengyan, 2009) developed an Improved Pareto Genetic Algorithm which is requires to select a best route between nodes while balancing between cost assignment and distance assignment of the route. This approach is considered as optimal when it achieves the less cost of the whole travel. TSP can be applied also in the iron and steel industry, where orders are scheduled on the hot rolling mill in such a way that the total set-up cost during the production can be minimized. The details of a recent application of modeling such problem are presented in (Tang, 2000). Here, the orders are treated as cities and the distance between two cities is taken as penalty cost for production changeover between two orders. The solution of the model will yield a complete schedule for the hot strip rolling mill.
(Chen, 2015) developed a new algorithm called Estimation of Distribution Algorithms (EDA) to solve the multiple travelling salesmen problem. This approach is used to minimize the total travelling distance and to maximize the distance. It is based on self-guided genetic algorithm with a minimum loading assignment (MLA) rule. The researcher found that his approach did not cause a longer travelling distance when the number of salesmen is increased from 3 to 10. (Crisan, Pentea and Palade, 2017) used the Romanian TSP instance with the main human settlements in order to derive several sequences of instances. This approach is applied to build new instances of TSP using geographic information systems features. These instances are then modified recursively using Lin–Kernighan method which shows to produce more stable instances.
Commuity health workers can be modeled as a TSP as per (snyders and Lane-Vissers, 2012). In this model the works have to visit the patients to collect data. They proposed an approach to by optimizing route planning. One of the techniques developed to solve multiple TSP is to use GA with multi-chromosome individual representation, where each salesman is assigned a separate chromosome. This approach which is developed by (Király and Abonyi, 2014), is integrated into Google Map to visualize the supply structure of the mobile mechanics in order to select right materials and delivered to the right place within right time
This paper proposes a prioritized multi-objective TSP (PMOTSP). The idea of the problem is as follows:
As an application of TSP, a salesman is requested to deliver the goods to customers in different locations. However, the order of visiting the customers' locations is predefined by a priority matrix. The proposed approach studies the relation between delivering goods to customers using shortest path so that it is beneficial for the service provider, and delivering it based on customers' priority so it is beneficial for the service receiver. In this paper, a multi-objective TSP is proposed to balance between shortest path and high priority. This paper applies genetic algorithm as a foundation to design the genetic optimization algorithm.

III.THE MATHEMATICAL MODEL OF PMOTSP

Suppose that there are a set of cities numbered 1, 2, ..n. The distance between two cities i and j is dist[i],[j] such that i≠ j. The order of visiting these cities is defined in a matrix form where the entry $M_{i,j}$ in this matrix represents the priority of visiting city j from city i and is defined as per[i][j], i≠ j.. The mathematical model of PMOTSP can be expressed as follows:

$$\begin{cases} f1(x) = \sum_{i,j \in x} dist[i],[j] \\ f2(x) = \sum_{i,j \in x} per[i],[j] \end{cases} \qquad (2)$$

Where x represents a route permutation of n cities. So function f1 is the summation of the distance from one city to the next city according the sequence defined in x, and function f 2 is the summation of the priorities from one city to the next.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

IV. THE PROPOSED GA FOR SOLVING PMOTSP

GA is adopted in this paper to solve the PMOTSP. The key feature of the proposed GA lays in the fitness function that is developed specially for this model. The description of the model will start by describing the chromosome representation for the proposed model.

A. Chromosome Representation.

Chromosomes in GA are used to represent the possible solution for the problem under consideration. The chromosome consists of genes that when integrated together forms the complete solution. For TSP, the adopted representation is to have integer number for each city and use these numbers when referencing these cities. Each chromosome consists of a sequence of integers to represent a possible path where the salesperson is visiting. The order of these cities within the chromosome forms the order of cities to be visited. A problem with 7 cities is represented as a chromosome of 7 genes as shown in Figure 1. This chromosome forms a possible solution that is a path starts from city number 2 passes through cities 5 - 1 -4- 6- 7 -3 and ends at 2 again, since the path must be closed.

2	5	1	4	6	7	3

Figure 1: sample chromosome for a problem of 7 cities

B. Initial Generation Creation

The initial generation consists of a set of chromosomes. The genes of these chromosomes are selected randomly from the search space. This search space consists of all cities. Once the city is selected, it cannot be selected again, which means that the city will be visited only once. Moreover, all cities need to be selected during the formation of the chromosome. Hence, the length of the chromosome is equal to the number of cities in the data set instance.

C. Parent selection

Once the first generation is created, the operators of GA are applied. The first operator of GA is parent selection. The purpose of parent select operator is to select two good individuals to perform crossover.

The most popular selection technique is simple random sampling selection which also called proportional selection. It has been applied by many researches [16-23]. This method performs roulette-wheel selection, where each individual is represented by a space that proportionally corresponds to its fitness. Stochastic sampling is used to choose individuals by repeatedly spinning the roulette wheel. This method may speed up the convergence with small fraction and avoids early premature convergence since good individuals have high probability of being selected for crossover.

In tournament selection [24], a group of i individuals are randomly chosen from the population. This group takes part in a tournament and an individual with highest fitness value wins. If i is chosen to be two, this method is called binary tournament selection. To further enhance this selection, i is selected to be five so each time from each 5 individuals, the best is selected.

In order to speed up convergence and avoid creating large number of generations, elitism technique can be applied as a selection technique [25]. In this technique best l members from the current generation are selected to form the mating pole for next generation. It is applied to ensure gradual improvement of the solution. In this paper, l is selected to have best two thirds of the current generation, i.e. the best 67% of the current population to form a mate for the parent selection. However, a validation process, before applying crossover, is applied to insure that the two selected parents are different from each other, i.e. $P1 \neq P2$.

D. Fitness function

The performance of MOPTSP is evaluated using a fitness function. This function is developed by combining two parts. This first part is the summation of the distances between cities that form the proposed path. The distance is calculated as follows:

$$\mathrm{dist}(c_k, c_j) = \sqrt{(x_k - x_j)^2 + (y_k - y_j)^2} \qquad (3)$$

Where c_k is the city that is in k^{th} position of child chromosome and c_j is the neighbour city of c_k which may be on the right or left of city c_i, that is at location k+1 or k-1. Knowing that cities k and j have the coordinates (x_k, y_k) and (x_j, y_j). The summation of these distances is calculated as

$$d = \sum_{i=1}^{n-1} \mathrm{dis}(i, i+1) \qquad (4)$$

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

The second part of the fitness function evaluates the priority of visiting these cities. The formulation of the fitness functions is as follows:

$$p = \sum_{i=1}^{n-1} \operatorname{per}(i, i+1) \qquad (4)$$

Where per(i, i+1) is the priority of visiting city (i+1) from city i. The priority matrix of visiting city (i+1) from city i is constructed randomly at the beginning of executing the program since there is no similar data is available in the literature.
In order to determine the k_{th}-gene of an offspring chromosome, a heuristic function ρ which is defined as the inverse distance between the cities, will be used to calculate the fitness of two pairs of cites [2]. This function is defined as:

$$\rho = \frac{1}{\operatorname{dist}(c_k, c_j)} \qquad (5)$$

The chromosome performance is measured using the fitness function that consists of two components. The first component is the distance weight. The second component is the priority weight. The function is as follows:

$$f = \alpha \cdot d + \beta \cdot p \qquad (6)$$

Where d is the total distance as calculated in (4), and p is the summation of priorities of visiting the cities forming the path. α and β are coefficients used to provide different weights for each component of the fitness function.

E. Crossover Operator
Crossover is considered as one of the main operators that heavily affects the performance of GA system. Researchers have proposed and analysed many techniques of crossover in order to study their effect on the output from several perspectives [26]. In all cases, crossover is performed such that the order of the genes, hence the developed path, is changed in a systematic manner to produce a new visible solution.
Among several crossover techniques, the one-point crossover is adopted. Although it may be not the best technique, but this open the floor for other researches to investigate the effect of other crossover techniques on the performance. The crossover probability is set to one in this research. That means that every parent selection ends by a crossover process.

V. EXPERIMENTS AND RESULTS

A. Data set description
In order to investigate the performance of the proposed techniques, several experiments need to be conducted on a benchmark data set. The data set adopted for these experiments is the well known TSPLIB data set (Reinelt, 1991), in which 8 instances of it are used here. These instances range in size from 51 cities to 150 cities. However, this data set consists of collections of cities along with the path length between each two cities. In order to companion the priority of visiting each city, a new matrix is constructed randomly for each data set instance to include such information.
The proposed GA system starts by creating initial generation consisting of 50 chromosomes, which is the population size. Recall that each chromosome forms a possible solution (travelling path). Hence the length of the chromosome is equal to the number of cities of the applied data set instance.
The stopping criterion of the proposed system is to have no more enhancement of the generations. The crossover probability is chosen to be 1. The stopping criteria for MOPTSP is one of two options; either the number of generations reaches 20 (this small number is chosen based on the experiments performed) or the difference in performance between two consecutive generations is less than a predefined threshold.
The parameters of the proposed GA are listed in Table 1.

TABLE 1 PARAMETER SETTING OF GABIR

Parameter Description	Value
Population size	50
Maximum number of generations	20
Chromosome length	No. of cities of the running set
Crossover rate	1
The number of best individuals forming the crossover mate (Elitism)	2/3 of the population
Mutation rate	0.1

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

B. Analysis of the results

The aim of these experiments is to find a proper value for the coefficients α and β, that is to balance between the priority of visiting a city and the cost of visiting it in terms of path length. For that, several experiments are conducted for each selected instance of the data set. The results considered in the analysis are those obtained from the last generation. Three different values for both α and β of formula (6) are investigated: The first one is $\alpha=0.5$ and $\beta=0.5$. This combination is referred as Case 1. These values give equal weight for both factors of the fitness function, which are the path length and the priority of visiting the cities. The second combination or Case 2 is to use $\alpha=0.75$ and $\beta=0.25$. These values give high importance for the path length. The last combination or Case 3 is to use $\alpha=0.25$ and $\beta=0.75$, which gives high importance to the priority of visiting the city. The optimal solution is the one that has a maximum score obtained by the fitness function out of these three cases.

In order to visualize the results properly, data instances 100A, 100B, 100C and 130 are normalized.

The results are listed in Table 2 and depicted in Figure 2. From Table 2 it can be seen that the highest fitness value is achieved when both α and β have the same value, that is Case 1. In another words, when both the path length and the priority of visiting the cities have the same weight or given the same importance. However, the performance of case 1 is still better than that of Case2. The fitness gets worse (higher) when the importance is given to the path length rather than the priority of visiting the cities, i.e. Case 2. While the smallest fitness value is achieved in the Case 3 with $\alpha=0.25$ and $\beta=0.75$. This implies that the best solution is achieved when the high weight or the importance is assigned to visiting the cities while the low weight or the low importance is designated to the length of the path. This result is achieved by all considered instances of the data set which means that this assignment of coefficients is consistent. These results are similar for each instance of the data set.

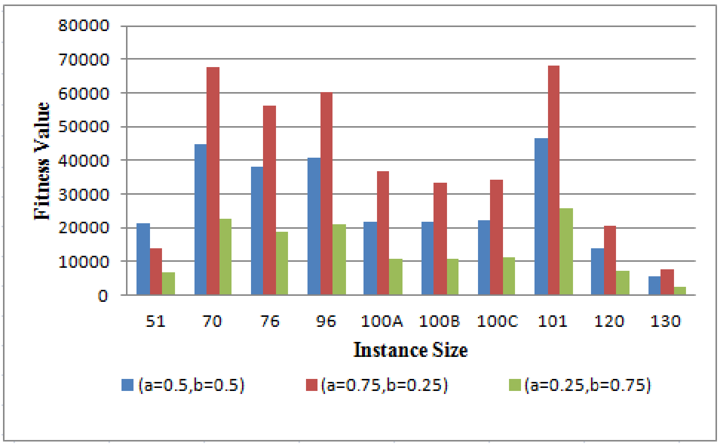

Figure 2: The performance of 10 instances for all cases

VI. CONCLUSION

TSP is one of the NP optimization problems, in which GA is considered as an efficient techniques used to solve it. In this paper a multi-objective TSP is considered. These objectives are finding a shortest path that needs to be used to visit all cities, and the priority of visiting the cities.

TABLE 2 THE FITNESS OF EACH DATE SET FOR DIFFERENT α AND β VALUES

Size	Case 1	Case 2	Case 3
	(a=0.5,b=0.5)	(a=0.75,b=0.25)	(a=0.25,b=0.75)
51	21535	13728	6723
70	44909	67492	22619
76	37983	56195	18697
96	40949	59990	20975
100A	21785	36989	1090760
100B	21678	33064	1068724
100C	22341	34014	1107226
101	46294	68316	25953
120	13695	20403	73072
130	53658	78510	26050

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

A matrix of priorities is created in addition to 10 instances of the well known TSPLIB are adopted to investigate the performance of this approach. The main part of GA in the current approach is the fitness function that combines two factors in one formula. These factors are the path length and the priority of visiting the cities. Each factor is associated a weighting coefficient. The experiments are conducted to determine the best values for these coefficients that lead to the highest performance. The results show that the smallest fitness value which means the best performance is achieved when the priority is assigned 75% and the path length is assigned 25%.

To generalize the results and further demonstrate its efficiency, several experiments need to be conducted, such as those to investigate other values of the coefficient, need to apply the approach on larger instance sizes, and to analyse the time performance.

REFERENCES

[1] J. Jing and M. Lidong, “The Strategy of Improving Convergence of Genetic Algorithm’. TELKOMNIKA, Vol.10, No.8, December 2012, pp. 2063~2068
[2] Aimin Zhou, and Feng Gao, 2013. A Decomposition Based Estimation Of Distribution Algorithm For Multiobjective Traveling Salesman Problems. 1st ed. Shanghai.
[3] Applegate, D., Bixby, R., Chvátal, V., Cook, W., 2003. Implementing the Dantzig-Fulkerson-Johnson algorithm for large traveling salesman problems. Mathematical programming. 97(1-2), 91-153
[4] Bryant, Kylie. 1st ed. 2015 Genetic Algorithms And The Traveling Salesman Problem..
[5] Christofides, N., Vehicle Routing. In: Lawer, E. L., Lenstra, J. K., Rinnooy Kan,A. H. G. and Shmoys, D. B. (eds.) The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, 431–448. New York: John Wiley & Sons, 1985.
[6] Cs.wmich.edu,. 'Department Of Computer Science | Western Michigan University'. N.p., 2015. Web. 10 May 2015.
[7] Euler, L., Solution d’une Question Curieuse Qui ne Parait Soumise a Aucune Analyse, Berlin 15, pp. 310–337, 1759, Published in: Opera Omnia, Vol. 1, No. 7, pp. 26–56, 1766.
[8] Flood, M. M., (1956), The Traveling Salesman Problem, Operations Research 4(1), 61–75.
[9] Garfinkel, R. S., Motivation and Modeling. In: Lawer, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G. and Shmoys, D. B. (eds.) The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, 17–36. New York: John Wiley & Sons, 1985.
[10] Gerard Reinelt, 1994. The Traveling Salesman: Computational Solutions for TSP Applications. Springer-Verlag, 1994.
[11] Google.com.bh,. 'Traveling Salesman Problems'. N.p., 2015. Web. 1 May 2015.
[12] Gupta, Namit, and Rajeev Kumar, 2012. Using New Variation Crossover Operator Of Genetic Algorithm For Solving The Traveling Salesmen Problem In E-Governance. 5th ed. New York Science Journal.
[13] Johnson D.S. & McGeoch L.A., 1995. The Traveling Salesman Problem: A Case Study in Local Optimization, November 20, 1995.
[14] Klan,ek, Uro‹.2011 Using The TS P S Olution For Optimal Route S Cheduling In Construction Management. 1st ed. 2011.
[15] Laporte, G. & Nobert, Y., 1980. A cutting planes algorithm for the m-salesmen problem. Journal of the Operational Research Society, Vol. 31, pp.1017–23.
[16] Matai, Rajesh, and Surya Surya Prakash Singh. Traveling Salesman Problem: An Overview Of Applications, Formulations, And Solution Approaches. 1st ed. 2010. Web. 3 May 2015.
[17] Menger, K., 1932, Eine Neue Definition der Bogenlange, Ergebnisse Eines Mathematischen Kolloquiums, Vol. 2, pp. 11–12.
[18] Renaud, J., Boctor, F.F., Ouenniche, J., 2000. A heuristic for the pickup and delivery traveling salesman problem. Computers and operations research, 27(9), 905-916.
[19] Roodbergen, K.J. and De Koster, R. 2001, Routing order pickers in a warehouse with a middle aisle. European Journal of Operational Research 133(1), 32-43.
[20] Shi Lianshuan, Li Zengyan, 2009, "An Improved Pareto Genetic Algorithm for Multi-objective TSP," Natural Computation, 2009. ICNC '09. Fifth International Conference on, vol.4, no., pp.585-588, 14-16 Aug. 2009.
[21] Sites.google.com,. 'Step5: Travelling Salesman Problem - Mohammad Shidujaman, NTPU'. N.p., 2014. Web. 23 May 2015.
[22] Tang, L.; Liu, J.; Rong, A. & Yang, Z. 2000. A multiple traveling salesman problem model for hot rolling scheduling in Shangai Baoshan Iron & Steel Complex. Uropean Journal of Operational Research, Vol. 124, pp. 267–82.
[23] Vali, Masoumeh. 1st ed. Isfahan: N.p., 2015. Solving Traveling Salesman Problem By Marker Method.
[24] Yeh, J.-Y., Lin, J.-Y., Ke, H.-R., and Yang, W.-P. (2007). Learning to rank for information retrieval using genetic programming. In Proceedings of ACM SIGIR 2007 Workshop on Learning to Rank for Information Retrieval (LR4IR '07), pp. 41-48. Amsterdam, Netherlands.
[25] Asllani, A. and Lari, A. (2007). Using genetic algorithm for dynamic and multiple criteria web-site optimizations. European Journal of Operational Research, vol. 176, no.3, pp. 1767-1777.