STATISTICAL COMPARISON OF ARTIFICIAL INTELLIGENCE TECHNIQUES FOR TRAVELING SALESMAN PROBLEM

[1]BURCU OZCAN, [2]YILDIZ SAHIN, [3]HILAL DINDAR, [4]ZEYNEP BIRCAN BAGDATLI

Kocaeli University
E-mail: [1]burcu.ozcan@kocaeli.edu.tr, [2]yildiz.sahin@kocaeli.edu.tr, [3]hilal_dindar@hotmail.com, [4]zebibalee@gmail.com

Abstract - There are many artificial intelligence techniques developed for solving optimization problems. Simulated Annealing and Genetic Algorithms are artificial intelligence methods and are still being developed. In this study, the Genetic Algorithms and Simulated Annealing methods are applied to the Traveling Salesman Problem which is a NP-hard optimization problem. The data was obtained from a regional distributor. For both methods, the objective function is determined to be the shortest total path. Significance of the results obtained by Simulated Annealing and Genetic Algorithms methods have been investigated statistically.

Index Terms - Artificial Intelligence, Genetic Algorithms, Simulated Annealing, Traveling Salesman Problem

I. INTRODUCTION

Artificial Intelligence Techniques are methods that have a significant share among different solution methods of optimization problems [1]. Artificial Intelligence, by far the most general definition, is a scientific discipline that aims both to simulate cognitive systems (rational / think like human) and to construct "intelligent" systems (rational / behave like human) [2].

Genetic Algorithms (GA) and Simulated Annealing Algorithms (SA), which are among Artificial Intelligence Techniques, can be used to solve a wide range of optimization problems. These methods produce problem-based solutions and do not give a definitive result. Their aim is to find good solutions.

In the first step of the GA optimization process, a good approximate solution is not needed as an initial step. In this first step, randomly generated sequences are used that contain relevant information about the problem [3].
The Simulated Annealing is a metaheuristic optimization method introduced by Kirkpatrick et al based on the simulation of thermal cooling of solids [4]. The method starts with a high temperature value. In each calculation step, numerous solutions are produced from the neighbors of the existing solution. These new solutions are either accepted or rejected according to the criteria specified. After each calculation step, the temperature is reduced according to a specified function. The algorithm is terminated either to the desired number of iterations, or to the minimum temperature value or when the desired solution is reached.
Admission of bad solutions is based on Boltzmann's probability distribution. In a system that is in thermal stability, the energy at temperature T(E) is distributed probabilistically according to equation (1). While the system can be in any energy state at high temperatures, it is very unlikely to be in high energy state as the temperature decreases [5].

P (E) = energy level, k = Boltzmann constant
P(E)=e^(-E/kt) (1)

When comparing the first solution and the new solution produced according to the neighborhood function, the fitness values are checked first. If the fitness value of the new solution is better than the previous solution ratio, the new solution is considered directly regardless of any other constraint. However, if the fitness value of the new solution is worse than the previous solution, a probability calculation is made using the following formula. The solution is accepted or rejected depending on the probability calculated according to the given formulation.

e- (|u2-u1|)/Tc (2)
In this formula, U2 is the fitness value of the new solution, U1 is the fitness value of the previous solution, and Tc is the temperature value at that moment [6] More detailed procedures of SA can be obtained from the literature [7, 8].

Genetic algorithms and simulated annealing, among the artificial intelligence techniques, are often used to find optimal solutions to complex problems. In this study, genetic algorithms and simulated annealing are briefly mentioned and both methods are applied on the traveling salesman problem. The Traveling Salesman Problem will be solved with two different methods and the best results will be determined. According to the obtained results, the methods will be statistically compared, and it will be determined which method gives better results. The MATLAB program was used to solve the problem.

II. TRAVELING SALESMAN PROBLEM

The Traveling Salesman Problem (TSP) is a combinational optimization problem that has been

extensively studied in the literature with a wide range of applications. For a given number of visit points, it is the problem of performing the return process starting from one point and visiting all the points once to minimize the total turn length. The Multiperiod Traveling Salesman Problem (MPTSP) is the same problem periodically designed. MPTSP is a different version of the Traveling Salesman Problem (TSP) and can be defined as a problem of visiting cities (customer points) in a certain time period (M-day). At least one visit is carried out every day within the specified M days. At the same time, all customers should be visited at the specified time [9].

The mathematical description of the Traveling Salesman Problem is as follows.

Minimize: $z = \sum_{i=1}^{n}\sum_{j=1,i\neq j}^{n} x(i,j)d(i,j)$
Criteria $\sum_{j=1,i\neq j}^{n} x(i,j) = 1, i = 1,2,\dots,n$
$\sum_{j=1,i\neq j}^{n} x(i,j) = 1, j = 1,2,\dots,n$
$\sum_{i,j\in S, i\neq j} x(i,j) \leq |S| - 1, \forall S \subset \{1,2,\dots,n\}$
$x(i,j) = \{1, if\ going\ from\ point\ i\ to\ point\ j\}$

$x(i,j) = \{0, if\ is\ not\ going\ from\ point\ i\ to\ point\ j\}$

III. SOLUTIONS FOR TRAVELLING SALESMAN PROBLEM

The data was obtained from a regional distributor. The province is divided into four parts for distribution purposes. In each of these regions, visits were determined based on experience, and the route of the vehicle to transport to the relevant region was established.

The routes in these regions, the latitude-longitude information of the stores on the routes, and the total tour distance of each route are obtained from the company. The firm's vehicles are able to meet the total amount of transportation to be made to the sales outlets in each region. The company wants better routes to be determined on a regional basis to reduce distribution costs. The problem of the firm is in accordance with the traveling salesman problem examined in this study. The problem used in this study is taken from the literature, the problem size is 29x29 and the source is mentioned in the bibliography [10].

In the solution of the problem, the population size 100 in the genetic algorithm and the parameters used in simulated annealing T = 10000 Tend = 0.1 Cooling coefficient = 0.9 were accepted.

The genetic algorithm and simulated annealing solutions of the above traveling salesman problem are as follows. The study was conducted in 100 iterations. All the results obtained are shown in Table 1.

Table 1: Results obtained using MATLAB

Simulated Annealing Results				Genetic Algorithm Results			
3603	3624	3539	2735	3123	3399	2419	2632
3466	3644	3584	3141	3075	3100	3060	2698
2500	3338	3725	2846	2404	2677	2475	2593
2434	3566	3386	3047	2839	3010	2564	2566
2819	3425	3738	2434	2418	2768	2946	2593
3957	3736	3773	2348	3034	2985	2912	2907
3873	3622	3848	2930	2566	2864	2574	2396
3914	2948	3499	2755	2735	3002	2799	2295
2660	3094	3504	2910	3115	2348	2735	2706
2896	3621	3501	3665	2678	2424	3141	3210
3842	3603	3604	3480	2903	2613	2846	3012
3991	3965	3592	2562	2539	2708	3047	2776
3700	2665	2907	2295	2471	2878	2434	2423
3655	3752	2295	2750	2813	2447	2348	2434
3451	3451	2348	3224	2676	2465	2930	2744
3803	3940	2562	2562	2500	2656	2755	2400
3572	3483	3224	2295	2434	2847	2654	2402
3676	3527	3127	2593	2614	3270	3480	2910
2539	3675	2475	2698	2819	2673	2764	2676
3603	3862	3060	2750	2896	2503	2562	2746
3684	3608	2731	2295	2479	2641	3224	2716
3950	3504	2562	2716	2603	2907	2572	2636
3719	3567	2993	2396	2660	2833	3127	2679
2676	3812	2574	3210	2948	2534	2993	2833
3903	3622	3127	2434	2665	2731	2750	2947
MIN			2295	MIN			2295
MAX			3991	MAX			3480
MEAN			3228,89	MEAN			2743,31

When the results are examined, it is seen that the two methods give the shortest distance as 2295 km. However, when evaluated in 100 iterations, the best value provided by the annealing simulation algorithm is 3991 km, which is far from the best result. When the averages of the results are compared, it was observed that the genetic algorithms produced better results than 3228,89 km obtained from simulated annealing with 2743,31 km.

In order to compare the methods statistically, some tests have to be done. These tests are parametric or nonparametric tests according to the structure of the data. To use parametric tests, the data must comply with certain assumptions. The data must be normally distributed, and the variances of the data must be homogeneous.

Table 2: Normality Test

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Annealing	,175	100	,000	,911	100	,000
Genetic	,071	100	,200*	,973	100	,039

Whether the data have normal distribution or not were checked via normality test in SPSS and the results are as shown in Table 2. The normal distribution matching of the values obtained from the methods is shown in Figure 1 and Figure 2.

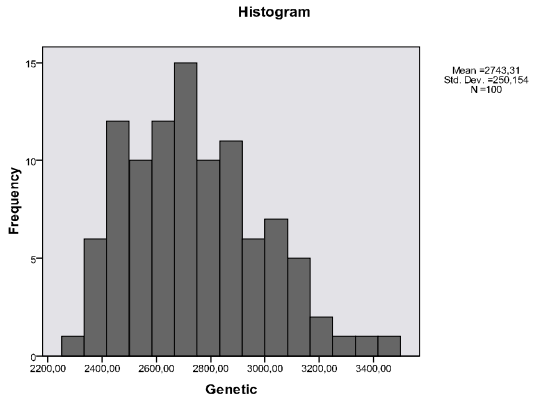

Fig 1: Genetic Algorithm Histogram Graph

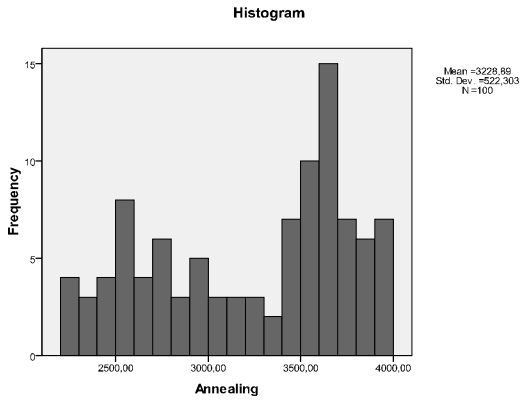

Fig 2: Simulated Annealing Histogram Graph

The Kolmogorov-Smirnov test results should be considered as the sample size in the data set is greater than 50. Since Sigma Value is less than 0.05, the data do not seem to fit the normal distribution. For this reason, it can be concluded that the use of non-parametric tests is more appropriate [11].

Whether or not there is a statistical difference between the simulated annealing and the genetic algorithm methods used to solve the traveling salesman problem can be found by using the Mann Whitney U test because binary comparisons and nonparametric methods are used.

Hypothesis is established and tested.

H_0: The methods are not statistically different from each other.

H_1: The methods are statistically different from each other.

Table 3: Mann-Whitney U and Wilcoxon test

Test Statistics			
Mann-Whitney U	2425,000	Wilcoxon W	7475,000
Z	-6,292	Asymp. Sig. (2-tailed)	,000

When the test statistic value is considered, H_0 is rejected as the sigma value is smaller than 0.05 as shown in Table 3. In other words, these two methods differ statistically.

CONCLUSION

In this study, the solution of the Traveling Salesman Problem known as the NP-hard problem, has been realized by Simulated Annealing Method and Genetic Algorithms, which are Heuristic Methods. As the result of these solutions, different optimum routes and distances have been obtained. When the results are examined, it is seen that the minimum distance is the same in both solutions. However, when 100 iterations are examined, genetic algorithms generally produce better results than simulated annealing algorithms. This study was compared with the smaller size (12x12) study using data from the Turkish Air Force's new generation cargo planes (A400M). In this study, the scheduling and routing problem for distributions among the potential bases in Turkey were investigated [12]. When the two methods were compared for this study, it is seen that the genetic algorithm gave the better results. The two methods used in problem solving are statistically compared. The normality test was applied to 100 values obtained and it was observed that the values did not fit the normal distribution. For this reason, Mann Whitney U test was used for nonparametric data. For hypotheses H_0 and H_1; H_0 means that there is no difference between the methods and H_1 hypothesis means that the methods are different. Since the sigma value of the test result is less than 0.05, the H_0 hypothesis is rejected and it is understood that these two methods are statistically different from each other. As a result, when the two methods are statistically analyzed in terms of both small size problems and large size problems, the genetic algorithm has been shown to perform much better than simulated annealing.

REFERENCES

[1] S.P. Leo Kumar, "State of The Art-Intense Review on Artificial Intelligence Systems Application in Process Planning and Manufacturing", Engineering Applications of Artificial Intelligence, vol. 65, pp. 294-329, 2017.

[2] H.A. Simon, "Artificial intelligence: an empirical science", Artificial Science, vol.77, no.1, pp. 95-127, 1995.

[3] V. Togan and A. Daloglu, "Genetik algoritma ile iki boyutlu kafes sistemlerinin şekil boyut optimizasyonu", Teknik Dergi, vol. 17, no. 2, pp. 3809-3825, 2006.
[4] S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi, "Optimization by simulated annealing", Science, 220(4598), pp. 671–680, 1983.
[5] Ke-Lin Du and M.N.S. Swamy, "Search and Optimization by Metaheuristics, Techniques and Algorithms Inspired by Nature", XXI-434, ISBN: 978-3-319-41191-0, Birkhauser Basel, 2016.
[6] S. Akkaş, "Karesel atama probleminin tavlama benzetimi ve paralel programlama teknikleri kullanarak çözümü". Pamukkale Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Yüksek Lisans Tezi, 2016.
[7] E.G.Talbi, "Metaheuristic: from design to implementation", Hoboken, New Jersey: John Wiley & Son, Inc, 2009.
[8] B. Fahimnia, H. Davarzani and A. Eshragh, "Planning of complex supply chains: a performance comparison of three metaheuristic algorithms", Computers and Operations Research, vol. 89, pp. 241–252, 2018.
[9] A.J. Orman, and H.P. Williams, "A Survey of Different Integer Programming Formulations of the Travelling Salesman Problem", In: Kontoghiorghes E.J., Gatu C. (eds) Optimisation, Econometric and Financial Analysis. Advances in Computational Management Science, vol 9. Springer, Berlin, Heidelberg, 2007.
[10] S. Çolak, "Genetik algoritmalar yardımı ile gezgin satıcı probleminin çözümü üzerine bir uygulama", Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi, vol. 19, no. 3, pp. 423-438, 2010.
[11] G.W. Corder and D.I. Foreman, "Nonparametric Statistics: A Step-by-Step Approach", John Wiley & Sons, 2014.
[12] B. Ozcan, H. Dindar and Z. Bağdatlı, "Genetik Algoritmalar ve Tavlama Benzetiminde Gezgin Satıcı Probleminin Cozumu, Yontemlerin İstatistiksel Karşılaştırılması" ISEOS 2017, XVIII. International Symposium on Econometrics, Operations Research and Statistics, ISBN: 978-875-6983-98-0, 2017.

★★★