iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47264

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Vehicle Routing Using Genetic Algorithm and Machine Learning

Shreya Sunil[1], Aishwarya Pardeshi[2], Yash Kakde[3], Himaja Namala[4]
[1, 2, 3, 4]*Dept. of Computer Engineering AISSMS Institute of Information Technology, Pune, India*

Abstract: ***Vehicle routing is the design and assignment of routes to vehicles according to different constraints such as traveltime, length of route, etc.The objective of the vehicle routing isto minimize the total route cost. Because you want to minimize costs spent on traveling, you want to find out the most efficient route, one that will require the least amount of traveling.***
Our goal is to use machine learning to forecast travel times between each two locations and then use the genetic algorithmto find the best route for a person to travel. We can apply machine learning to anticipate journey times between each two sites and the genetic algorithm to discover the ideal deliverytruck travel schedule. Genetic algorithms (GA) are an intriguing way to solve issues involving search and optimization. For more complex situations, little has been done to provide a step-by-step implementation of a GA in Python.
Index Terms: Genetic Algorithms(GA), Traveling SalesmanProblem(TSP), Vehicle Routing(VR)

I. INTRODUCTION

The ”Traveling Salesman Problem” is abbreviated as TSP. It’s an NP-hard combinatorial problem. The problem involves a salesperson and a number of communities. Starting from a certain location (for example, a head office), the salesman must travel across all of the cities before returning to the starting city. The purpose of this task is to lower the overall distance travelled by the salesman. The following is a mathematical representation of the travelling salesman problem.
The Traveling Salesman Problem is solved using a genetic algorithm in this research paper. A genetic algorithm is a method for estimating computer models that is based on approaches derived from the area of biology’s genetics. To employ this method, one must first encode potential model behaviours into ”genes.” The current models are graded and permitted to mate and reproduce based on their fitness after each generation. The genes are transferred during mating, and crossings and mutations can occur. The present population is eliminated, and the kids of the current population become the next generation. Genetic Algorithm also encompasses a wide range of modelling and optimization approaches. The thing being modelled is usually represented in a way that is simple to change automatically. The data is then used to produce a huge number of candidate models, which are then evaluated against it. The ”best” models are kept for the following generation once each model is assessed. The procedure is then repeated until the models are in agreement (as in asexual reproduction). If the model is created in such a manner that the victors have ”genes,” they can ”mate” to generate the following generation. These methods are examined in conjunction with the Travel- ing Salesman Problem(TSP). The Problem’s goal is to identify the quickest route for a travelling salesperson who must start from his home city and visit each location on a list exactly once before returning home. The biggest challenge with this topic is the enormous number of potential tours: (n-2)! /2 for n cities. Probabilistic search methods that replicate natural evolution are known as evolutionary algorithms.

II. LITERATURE SURVEY

In the genetic algorithm, there are many parameters needing to be set in advance. To figure out the effects of parameters in GA, Chutian Sun et al. [1] conducted experiments with different parameters and compared the performance.
Anuj Kumar et al. [2] In this work for population initial- ization the author applied the greedy algorithm in Genetic Algorithm.In this greedy permutation approach applied on different problems which are a major concern in the Traveling Salesman Problem and find the better results as compared with an existing method, NF and this system make genetic algorithms much more reliable and efficient than methods which based on random initialization.
Heuristic methods and genetic approaches are the most appropriate ways to solve the travelling salesman problem. Multiple optimal solutions can be obtained for this problem by using various combinations of selection, crossover and mutation techniques. This was found out by Sahib Singh

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Juneja et al. [3] who surveyed many approaches and many combinations of genetic operators for this problem and the used combination of genetic operators in this paper is theoptimized one among them.

Usually, Traveling Salesman Problem has one element to beevaluated, distance. But in a real condition, other elements areneeded to be evaluated like the obstacles. This was done by IdaBagus Ketut Widiartha et al. [4] by classifying the obstacles into 2 categories: fixed and non-fixed. They found that the solution of TSP increases the distance between location, but minimizes the number of the obstacle.

Dr.Sabry M. Abdel-Moetty et al. [5] proposed an algorithm that constructs an offspring from a pair of parents using better edges on the basis of their values that may be present in the parents' structure maintaining the sequence of nodes in the parent chromosomes.

J. Wang et al. [6] proposed a multi offspring genetic algorithm and how is applicable in solving the Traveling Salesman Problem which we studied and gained many insightsfrom it.

A new genetic cross over operator using greedy approach was proposed by Vinod Jain et al. [7]. The algorithm has been implemented and results are better than existing cross over techniques. We are trying to include it in our method as well.

Dantzig, G. et al. [8] gave us more insights about the Traveling Salesman problem and we found answers to manyof our queries related to the large scale traveling salesman problem.

Ivan BrezinaJr. et al. [9] studied TSP using ant colony opti- mization and concluded that the quality of solutions depends on the number of ants. The lower number of ants allows the individual to change the path much faster. It works better withlower number of iterations.

III. PROPOSED SYSTEM

To optimize travel routes for a delivery vehicle by using machine learning model predictions. This is a two-component problem: first, train a machine learning model on the data to predict how long it will take a delivery vehicle to go from point one point to another, and then feed these predictionsinto a genetic algorithm which decides which is the most time efficient visit order for a given set of points.

A. Genetic Algorithm

1) Initialize the population randomly.
2) Determine the fitness of the chromosome.
3) Until done repeat:
4) Select parents.
5) Perform crossover and mutation.
6) Calculate the fitness of the new population.
7) Append it to the gene pool.

B. Genetic Operators

The genetic operator in genetic algorithms refers to a tool that guides the algorithm towards a solution to a given problem. They provide the basic search mechanism of the GA.

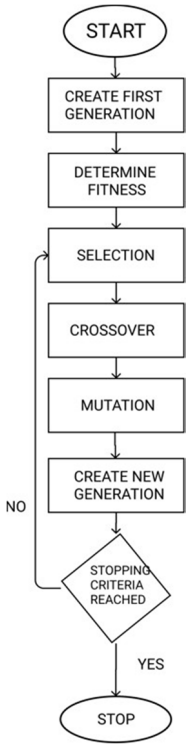
There are three main types of genetic operators: mutation(creation and maintenance of genetic diversity), crossover (combination of existing solutions into new solutions) and se- lection (choice between solutions), all of which work together in harmony to give the algorithm the best chance of success

1) *Selection:* selecting parents with highest fitness to create next generation. we're yet to decide which approach we'll be using, either roulette wheel selection or tournament selection or elitism. we'll perform selection and create mating pool.
2) *Crossover:* As we need to visit each location exactly once we'll be using ordered crossover which is a special breeding function. In ordered crossover, randomly select subset of first parent string and fill the remainder of the route with genes from the second parent in order, withoutduplication.
3) *Mutation:* We cannot actually drop any cities so we'll be using swap mutation wherein two cities will swap places in our route.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

IV. GENETIC ALGORITHM MODEL FLOW



V. METHODOLOGY

A. Data Selection

The very first step is selection of data. This means to gather the information from multiple sources to get an accuratepicture of the problem. We are using a Kaggle data set having the data set of taxis in New York.

B. Data Preparation

It is the second step which means to transform the raw data so that it can run through machine learning algorithms to makeaccurate predictions. The dataset provided by Kaggle is well- processed. The given dataset has each location's coordinates. By using this, we'll calculate Manhattan distance to get a senseof direction.

C. Data Analysis

In this phase we analyse the data visually by various visualization tools like Power BI, Tableau , etc.

D. Selecting the right model

As the data and goals are fixed and clear, we need to work with the right model which will give us the desired results. A low variance model is required and we want to add the direction of the route in the form of positive or negative numeric value. A boosted tree is preferred as they work very well with the given data set and they handle good amountof complexity. In addition to that LightGBM is also a good model to work with but when compared to XGboost it's a bit cumbersome to deal with.

E. Training and Prediction

In this step, the data that we are working with will be used to incredibly improve the model's ability to predict. Y=m*x+bY=output M=slope X=input B=Y-intercept As the trainingprogresses, we are closer to an ideal output.

©IJRASET: All Rights are Reserved

F. Evaluation

The final step is known as evaluation. It involves testingour model against an unknown data or a data that hasn't been used to train model. This helps us to determine the accuracyof our model.

VI. ALGORITHMS

A. Vehicle Routing Algorithm

An important application of optimization is vehicle routing,in which the goal is to find the shortest route for a fleet ofcars to travel between locations, usually by calculating the least amount of total distance or cost. The following are a few examples of routes optimized by optimization.

1) A package delivery company wants to assign routes fordrivers to make deliveries.
2) A ride-sharing company wants to assign routes for driversto pick up and drop off passengers
3) A cable TV company wants to assign routes for techni-cians to make residential service calls.

The problem has attracted a lot of attention all over the worldfor two basic reasons:

a) The problem appears in a large number of practicalsituations
b) The problem is theoretically interesting and not at all easyto solve.

B. Traveling Salesman Problem

The travelling salesman problem is a famous optimizationproblem requiring the least total distant Hamiltonian cycle.The Goal of TSP is to minimize a distance of route or to findthe shortest path to visit all of the locations in the given list.
Mathematically traveling salesman problem is formulatedas given:
TSP = (G, s, t): G = (V, E)

where G is a complete graph, s is a function
$V \times V \rightarrow Z$, t Z, G is a graph containing the distance to travel along with the cost which is not greater than t.
There are several methods for solving TSP such as Shortest Path Algorithms, Simple Insertion Algorithms, Elastic Net Methods, Simulating Annealing Algorithms, Ant Algorithms, and Genetic Algorithms. But the popular one among these methods is Genetic Algorithm (GA).

C. Genetic Algorithm

Evolutionary is the idea of GA. The method tries to mimic the process of natural selection. Every population of GA has chromosomes, where each of chromosome contains the infor- mation about the solution. Each of chromosome will evaluate using fitness function. The best fitness function will keep in every generation, and in the last generation chromosome with the best fitness function will be a solution of a problem. Selection finds a pair of offspring that will be used on the crossover. We need mutation to create another opportunityto take out the solution from local optima. Crossover andmutation will create another solution that will replace the worstsolution in every generation.

D. LightGBM

1) LightGB stands for Light Gradient Boosting Machine.
2) LightGBM uses Gradient-based One Side Sampling Technique.
3) LightGBM splits the tree leaf-smart rather than other boosting algorithms that develop tree stage-smart. It chooses the leaf with maximum delta loss to grow. Because leaf is constant, the leaf-smart algorithm has decreased loss in comparison to the extent-smart set of rules. Leaf-smart tree increase would possibly increase the complexity of the version and might lead to overfitting in small datasets.
4) After trying out LightGBM model on our data we found out the mean absolute error to be 4.956483768920135.

E. XGBoost

1) XgBoost stands for Extreme Gradient Boosting. It is a library which can be used to access a number of interfaces.
2) XGBoost is an enactment of gradient boosted decision trees which are mainly made for improved speed and per- formance. It optimizes the training for Gradient Boosting.

3) In gradient boosting, each predictor corrects its predeces- sor's error. In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost.
4) Weights are assigned to all the independent variables which might be then fed into the selection tree which predicts outcomes. The load of variables anticipated in- correct by the tree is extended and these variables are then fed to the second selection tree.
5) These individual classifiers/predictors then ensemble to offer a sturdy and extra particular version. It is able to paintings on regression, class, ranking, and user-described prediction troubles.
6) Features of XGBoost:
a) Gradient Boosting
b) Stochastic Gradient Boosting
c) Regularized Gradient Boosting
d) Parallelization of trees and using all CPU cores
e) Distributed Computing
f) Out-of-Core Computing for large datasets
g) Cache Optimization
h) Sparse Aware for working with missing values
i) Block Structure
j) Continued Training
7) After trying out the XGBoost model on our data we foundout the mean absolute error to be 4.827394962310791 which is lesser than that of LightGBM and hence we decided to go with XGBoost.

VII. IMPLEMENTATION

A. Data Selection

The very first step is selection of data. This means to gather the information from multiple sources to get an accuratepicture of the problem. We are using a Kaggle data set having the data set of taxis in New York.

B. Data Preparation

It is the second step which means to transform the rawdata so that it can run through machine learning algo- rithms to make accurate predictions. The dataset providedby Kaggle is well-processed. The given dataset has each location's coordinates..After that,we've converted datetime strings into datetime.Then we'll construct other variables like month,date etc.After that,we'll get latitude and longitude dif- ferences.Then, we'll convert duration to minutes for easier interpretation.Finally, we'll convert trip distance from the ob- tained longitude and latitude differences to Manhattan distanceto get a sense of direction

C. Data Analysis

In this phase we analyse the data visually by various visualization tools like Power BI, Tableau , etc.

D. Selecting the Right Model

As the data and goals are fixed and clear, we need to work with the right model which will give us the desired results. A low variance model is required and we want to add the direction of the route in the form of positive or negative numeric value. A boosted tree is preferred as they work very well with the given data set and they handle good amountof complexity. In addition to that LightGBM is also a good model to work with but when compared to XGboost it's a bit cumbersome to deal with.

E. Training Prediction

In this step, the data that we are working with will be used to incredibly improve the model's ability to predict. Y=m*x+bY=output M=slope X=input B=Y-intercept As the trainingprogresses, we are closer to an ideal output.

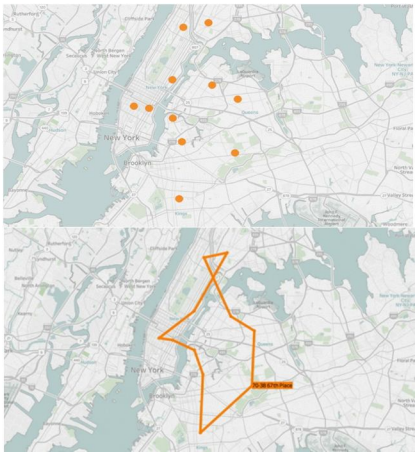
F. Evaluation

The final step is known as evaluation. It involves testing ourmodel against an unknown data or a data that has not been used to train model. This helps us to determine the accuracyof our model.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VIII. RESULTS

After putting both prediction models to the test,more basic XGBoost with no meteorological data performed somewhat better than LightGBM, with a mean absolute error of 4.8minutes vs 4.9 minutes for LightGBM. Used a genetic algorithm to build a sample solution and make a demo after selecting XGBoost and storing the model.



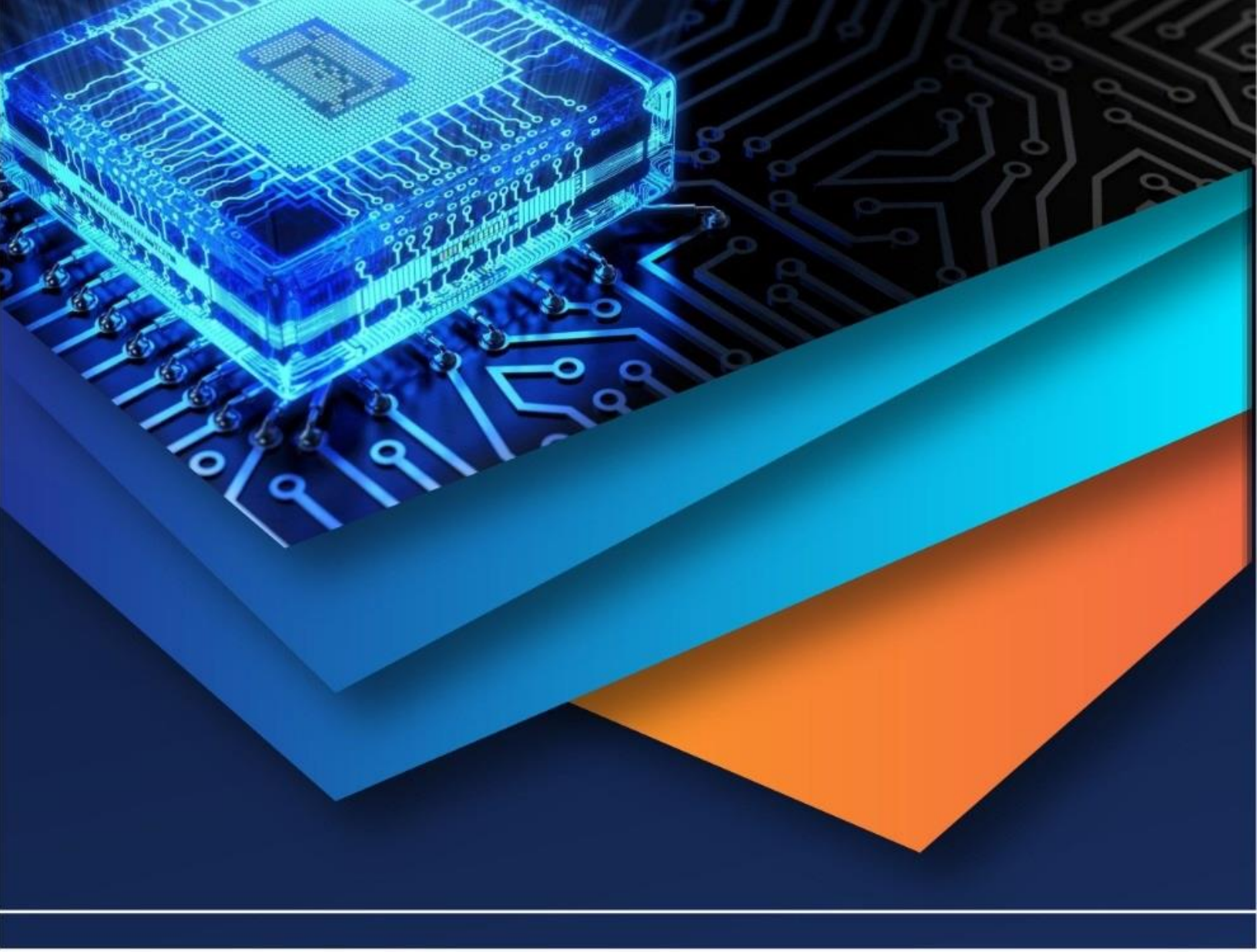
IX. CONCLUSION

Traveling salesman problem (TSP) is a combinatorial opti- mization problem. It is NP-hard, and the most intensely studiedproblem in the optimization area. But as the number of cities increases, the complexity of the problem increases. Using Genetic Algorithm approach, we have solved the Trav- eling Salesman Problem. The system starts with a matrix of the calculated Euclidean distances between cities to be visited by the travelling salesman and a randomly chosen order for the cities as an initial population.Then, new generations are created repeatedly until a stopping criterion is reached.

REFERENCES

[1] A Study of Solving Traveling Salesman Problem with Genetic AlgorithmChutian Sun Department of Industrial Engineering, Huazhong Universityof Science and Technology, Wuhan 430000, China.
[2] Improved Genetic Algorithm to Solve Small Scale Travelling Salesman Problem Anuj Kumar Department of Computer Engineering and Appli- cations GLA University Mathura, India
[3] Travelling Salesman Problem Optimization Using Genetic Algorithm Sahib Singh Juneja1, Pavi Saraswat2, Kshitij Singh3, Jatin Sharma4, Rana Majumdar5, Sunil Chowdhary6 1,2,3,4,5,6Amity School of Engi- neering and Technology, Amity University
[4] Traveling Salesman Problem Using Multi-Element Genetic Algorithm Ida Bagus Ketut Widiartha1, Sri Endang Anjarwani2, Fitri Bimantoro3 Informatics Engineering and Electrical Engineering Dept., Faculty of Engineering, Mataram University Mataram, Indonesia
[5] S. M. Abdel-moetty, “Enhanced Traveling Salesman Problem Solving using Genetic Algorithm Technique with modified Sequential Construc-tive Crossover Operator,” Int. J. Comput. Sci. Netw. Secur., vol. 12, no. 6, pp. 134–138, 2012.
[6] J. Wang, O. K. Ersoy, M. He, and F. Wang, “Multi-offspring genetic algorithm and its application to the traveling salesman problem,” Appl. Soft Comput. J., pp. 1–9, 2016.
[7] J. S. PRASAD AND V. JAIN, ”AN OPTIMIZED ALGORITHM FOR SOLVING TRAVELLING SALESMAN PROBLEM USING GREEDYCROSS OVER OPERATOR,” 2016 3RD INTERNATIONAL CONFER-ENCE ON COMPUTING FOR SUSTAINABLE GLOBAL DEVELOP-MENT (INDIACOM), NEW DELHI, 2016, PP. 2981-2984
[8] Dantzig, G., and Johnson, R. F. (1954) Solution of a large-scale traveling-salesman problem. Journal of the Operations Research Societyof America: 2(4), 393-410.
[9] Ivan BrezinaJr.,ZuzanaCickova, “Solving the Travelling Salesman Prob-lem using the Ant colony Optimization”, Management Information Systems, 2011, Vol. (6), No. (4).
[10] Naveen kumar, Karambir and Rajiv Kumar, “A Genetic Algorithm Approach To Study Travelling Salesman Problem”, Journal of Global Research in Computer Science, 2012, Vol. 3, No. (3).
[11] Davendra, D. 2010. Traveling salesman problem theory and ap- plicat ions. Rijeka, Croat ia: InTech. [Online]. Available at : ht tp://www.intechopen.com/books/t raveling-salesman-problemtheoryand-applicat ions, accessed on 15 April 2013.
[12] Xianxi Deng. Research and improvement of solving TSP problem with genetic algorithm Shenyang: Northeastern University of Information Science and engineering, 2008.
[13] Adewole Philip, AkinwaleAdioTaofiki and OtunbanowoKehinde, “A Ge- netic Algorithm for Solving Travelling Salesman Problem”, International Journal of Advanced Computer Science and Applications, 2011, Vol. (2),No. (1).
[14] Grefenstette, J., Gopal, R., Rosmaita, B., Van Gucht, D. (1985, July). Genetic algorithms for the traveling salesman problem. In Proceedings of the first International Conference on Genetic Algorithms and their Applications (Vol. 160, No. 168, pp. 160-168). Lawrence Erlbaum.

[15] Bennaceur, H. Alanzi, E. (2017). Genetic Algorithm For The Travelling Salesman Problem using Enhanced Sequential Constructive Crossover Operator. International Journal of Computer Science and Security (IJCSS), Volume (11) : Issue (3) : 2017.
[16] Yingying Yu, Yan Chen, Taoying Li. the improved genetic algorithm to solve the t raveling merchant problem [J]. Cont rol and decision, 2014, Vol. (29): 1483-1488.
[17] Milena Karova, Vassil Smarkov, St oyan P enev," Genet ic operat ors crossover and mutat ion in solving the TSP problem", Int ernat ional Conference on Computer Systems and Technologies - CompSysT ech' 2005.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)