

Optimization of Non-Linear Multiple Traveling Salesman Problem Using K-Means Clustering, Shrink Wrap Algorithm and Meta-Heuristics

R.Nallusamy¹*, K.Duraiswamy², R.Dhanalaksmi³, P. Parthiban⁴

¹Department of CSE, K.S.R.College of Technology Tiruchengode-637215, Tamilnadu, India

²K.S.R.College of Technology Tiruchengode-637215, Tamilnadu, India

³D-Link India Ltd, Bangalore, India

⁴Department of production engineering, National Institute of Technology, Tiruchirappalli, India

(Received 29 August 2009, accepted 27 September 2009)

Abstract: This paper deals with generating of an optimized route for multiple Traveling Salesman Problems. We used a methodology of clustering the given cities depending upon the number of salesmen and each cluster is allotted to a salesman. “*k*-Means clustering” algorithm has been used for easy clustering of the cities. In this way the *m*TSP has been converted into TSP which is simple in computation compared to *m*TSP. After clustering, an optimized route is generated for each salesman in his allotted cluster. To achieve this, we first generated a parent route using “Shrink Wrap” algorithm and this parent string is further optimized by using two other optimizing algorithms. For this purpose, “Tabu Search” and “Simulated Annealing” were extensively used. From the results, we observed that Simulated Annealing generate optimized route covering less distance than Tabu search.

Keywords: combinatorial optimization; *k*-means clustering; multiple traveling salesman problem; simulated annealing and Tabu search

1 Introduction

Problems of combinatorial optimization are characterized by their well-structured problem definition as well as by their huge number of action alternatives in practical application areas of reasonable size. Especially in areas like routing, task allocation, or scheduling such kinds of problems often occur. Their advantage lies in the easy understanding of their action alternatives and their objective function. Utilizing classical methods of Operations Research (OR) often fails due to the exponentially growing computational effort. Therefore, in practice heuristics and meta-heuristics are commonly used even if they are unable to guarantee an optimal solution.

1.1 Scope and objectives of research

A careful analysis of literature on the variants and methodologies of combinatorial optimization problems reveals the following: Some of the variants of combinatorial optimization problems are yet to be explored to solve using meta-heuristics techniques in the literature [9]. These include

- Multiple Traveling Salesman Problem (*m*TSP)
- *m*TSP with Balanced Allocation of Nodes with Single Objective, etc. . .

In TSP a number of cities have to be visited by a salesman who must return to the same city where he started. In solving the problem one tries to construct the route so that the total distance traveled is minimized. In the *m*-TSP problem, the *m*-salesman has to cover the given cities and each city must be visited by exactly

*Corresponding author. Mob.:+91-9865539172. E-mail address: nallsam@rediffmail.com

one salesman [11]. Every salesman starts from the same city, called depot, and must return at the end of his journey to this city again [1][2][4].

If n is the number of cities to be visited for the TSP then $(n - 1)!$ is the total number of possible routes. Following this basic formulation, an exponential relationship exists between the number of cities and possible routes, for instance if there are 5 cities there are 24 possible routes, for 6 cities 120, for 10 cities 362,880, and so on. As the amount of input data increases the problem increases in complexity, thus the computational time needed renders this method impractical for all but a smaller number of cities

The rest of the paper is organized as follows. The review of the existing work is given in section 2. Section 3 deals with the problem formulation. The clustering and optimization algorithm concepts are given in section 4. Section 5 deals with the implementation concepts and results of Simulated annealing and Tabu search. Conclusion and future scope is given in section 6.

2 Review of the existing work

Multiple travelling salesman problem handles more than one salesman. In m TSP, number of salesmen is more than one and salesman should be allotted few cities among the whole. This is done by dividing the cities into clusters and allotting each cluster to each salesman. Aristidis Likas; Nikos Vlassis, Jakob J. Verbeek (2002) suggested clustering of data points by using k means clustering. However, the resulting TSP is highly degenerate, when an m TSP is transformed to a single TSP since the resulting problem is more arduous to solve than an ordinary TSP with the same number of cities. Many methods have been suggested for obtaining optimized route [5][6][7][11].

A E Rizzoli, et.al. have focused on the Application of Ant Colony Optimization on the Vehicle Routing Problem and its real world application. They have worked on the aim to demonstrate the applications of ACO on different variants of the routing problem in the day to day activities [1]. Allan Larsen has worked on the dynamic factors of vehicle routing problem. In his thesis he has investigated the dynamics of the vehicle routing problem in order to improve the performances of existing algorithms and as well as developed new algorithms [2]. Various cluster analysis concepts have been discussed in [3].

Arthur.E.Carter, Cliff.T.Ragsdale have developed a new approach to solve multiple travelling salesman problem. The method proposes new set of chromosomes and related operators for the Multiple Traveling Salesman Problem (MTSP) and compares theoretical properties and computational performance of the proposed technique. The computational technique shows that the newer technique results in the smaller search space and produces better solutions [4]. Hannes Schabauer, et.al. have worked on to solve Traveling Salesman Problem(TSP) heuristically by the parallelization of self-organizing maps on cluster architectures [12].

Ding Chao, Cheng ye, He Miao have developed a two level genetic algorithm which favors neither intra-cluster path or inter-cluster path. The set of vertex data s which are given in the graph were portioned into clusters and now this problem becomes a cluster Travelling Salesman Problem (CTSP). The results from the paper indicate that the algorithm proposed is more effective than the existing algorithms. The variants that exist in the problem are properly dealt without affecting the constraints [10]. Klaus Meer has worked on the simulated annealing methodologies and has proved that this algorithm outperforms any metropolis or standard algorithm. And finally they have worked on the TSP to prove the above mentioned. The standard methodologies were adopted and the results were checked with Simulated Annealing [15]. Various shortest path routing concepts for networks were analysed in [8][13][14].

2.1 Research gap and proposed work

From the review, we understood that most of the problems involved solving the Conventional Traveling Salesman Problem using exact methods like Cut and Solve technique as well as meta-heuristic methods like simulated annealing and Genetic Algorithm for solving the same. They however scarcely dealt with the Multiple Traveling Salesman Problem which represents the realistic case of more than one salesman.

As far as the heuristic algorithms for the m TSP are concerned, the previous literature has an emphasis on artificial neural networks. To the best of our knowledge, from the literature review, no efficient meta-heuristic algorithms exist for the solution of large-scale m TSPs. Hence, an analysis is made and an heuristic

is formed to transform m TSP to TSP, and to optimize the tour of an individual. There was no specialized algorithm for the purpose of clustering (conversion of m TSP to TSP) or optimal allocation of cities to each salesman. Hence, we decided to deal with the less frequently approached and more realistic Multiple Traveling Salesman problem along with a specialized clustering heuristic, namely k-means clustering algorithm. This also breaks down a large sized problem into a simpler one. Optimization of route will be done by generating a parent route and optimizing the same using any relevant method. In our work, we have done the optimization by two methods. Then the effective solution generated by each method is studied and is compared with the results obtained from other methods.

3 Problem formulation

Due to the combinatorial complexity of the TSP, approximate or heuristic solution procedures are almost always employed in practice. The mathematical structure of the TSP is a graph where the cities are the nodes of the graph. Connections between pairs of cities are called edges and each edge has a cost associated with it which can be distance, time or other attribute. If n is the input number of vertices representing cities, for a weighted graph G , the TSP problem is to find the cycle of minimum costs that visit each of the vertices of G exactly once [6][7].

3.1 Multiple Traveling Salesman Problem

The m TSP can in general be defined as follows: Given a set of nodes, let there be m salesmen located at a single depot node. The remaining nodes (cities) that are to be visited are called intermediate nodes. Then, the m TSP consists of finding tours for all m salesmen, who all start and end at the depot, such that each intermediate node is visited exactly once and the total cost of visiting all nodes is minimized.

3.2 Mathematical Formulation of the TSP and Problem definition

There are many mathematical formulations for the TSP, employing a variety of constraints that enforce the requirements of the problem. Since this is not the appropriate forum for reviewing all of the potential formulations, one has been chosen in order to demonstrate how such a formulation is specified. The following notation is used:

- n – the number of cities to be visited; the number of nodes in the network
- i, j, k – indices of cities that can take integer values from 1 to n
- t – the time period, or step in the route between the cities
- x_{ijt} – 1 if the edge of the network from i to j is used in step t of the route and 0 otherwise
- d_{ij} – the distance or cost from city i to city j

The problem requires starting from a given city, visiting subsequent cities, and returning to the starting city. The optimal solution chooses the route that minimizes the total distance traveled. There are $(n - 1)!$ possible tours. The following is an example of one linear programming formulations of the TSP problem:

The objective function (Z) is to minimize the sum of all costs (distances) of all of the selected elements of the tour:

$$Z = \sum_{i=1}^n \sum_{j=i}^n \sum_{t=1}^n d_{ij} x_{ijt} \quad (1)$$

The tour is subject to the following constraints:

For all values of t , exactly one arc must be traversed, hence

$$\sum_i \sum_j x_{ijt} = 1 \text{ for all } t. \quad (2)$$

For all cities, there is just one other city which is being reached from it, at some time, hence

$$\sum_j \sum_t x_{ijt} = 1 \text{ for all } i. \quad (3)$$

For all cities, there is some other city from which it is being reached, at some time, hence

$$\sum_i \sum_t x_{ijt} = 1 \text{ for all } j. \quad (4)$$

When a city is reached at time t , it must be left at time $t + 1$, in order to exclude disconnected sub-tours that would otherwise meet all of the above constraints. These sub-tour elimination constraints are formulated as:

$$\sum_i x_{ijt} = \sum_k x_{jkt+1} \text{ for all } j \text{ and } t. \quad (5)$$

In addition to the above constraints the decision variables are constrained to be integer values in the range of 0 to 1:

$$0 \leq x_{ijt} \leq 1. \quad (6)$$

If the problem size increases, it is very difficult to solve in a reasonable polynomial time. We had developed all the below meta-heuristics techniques and compared for m TSP.

3.3 Assumptions

All the salespersons have to start from a common depot and after traveling through a set of cities, they should return back to the starting depot. There are no capacity constraints and no cost constraints. But, all the cities must be visited by any one of the salesperson and each salesperson has to visit a particular city exactly once.

4 Metaheuristics for route optimization

Efforts have concentrated on the development of heuristics that are not guaranteed to find the shortest tour, but are likely to quickly find either the optimal solution or a near-optimal alternative. They are approximate approaches based on algorithms that construct feasible solutions within reasonable computing time. Improvement heuristics are algorithms that start with an initial feasible solution and successively improve it through a sequence of exchanges.

4.1 Proposed heuristics

The solution to the problem is attained using three-stage heuristics. The first-stage involves the conversion of a m TSP to TSP using k -means Clustering algorithm. The second-stage of the heuristic comprises of forming the initial tour for m salesperson and is generated using Shrink-wrap algorithm. The third-stage is meta-heuristic approach comprising simulated annealing and Tabu search to optimize the tour for m salesperson. In this work we have taken 60 cities 6 salesmen problem for the route optimization purpose. Fig. 1 shows the coordinates of 60 cities.

4.2 Transformation of m TSP to TSP and calculation of shortest path

Clustering of the cities makes the calculations much simpler. The search space for the solution increases as the number of cities decreases and vice-versa. If there are N cities then the search space will be $N!$ and the computational time also high accordingly. Hence to reduce the burden of mathematical complexity N value should be reduced and this is achieved by clustering. The following heuristics were used for solving the given 60 cities 6 travelling salesmen problem. City number 33 is considered to be the headquarters of all the salesmen.

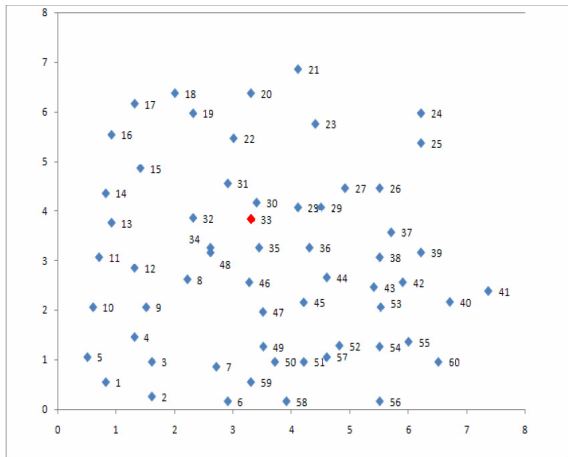


Figure 1: Plot of cities salesman has to travel

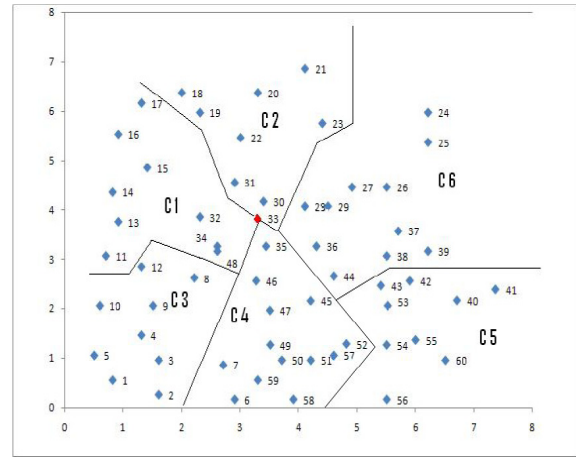


Figure 2: Clustered cities after k-means clustering

4.2.1 k-means Clustering

Simply speaking *k*-means clustering is an algorithm to classify or to group the objects based on attributes/features into *k* number of group. *k* is a positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid [3].

The algorithm is composed of the following steps:

- Place *K* points into the space represented by the objects that are being clustered. These points represent initial group centroids.
- Assign each object to the group that has the closest centroid.
- When all objects have been assigned, recalculate the positions of the *K* centroids.
- Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

Table 1 and Fig. 2 show the results, after applying *k*-means clustering.

Table 1: Final results of *k*-means clustering

Sales person	Cities Allotted	No. of cities
1	21-23-24-25-26-27-28-29-30-33-35-36	12
2	1-2-3-4-5-8-9-10-11-12-33-34-48	13
3	13-14-15-16-17-18-19-20-22-31-32-33	12
4	6-7-33-45-46-47-49-50-51-57-58-59	12
5	33-52-54-55-56-60	6
6	33-37-38-39-40-41-42-43-44-53	10

Table 2: Final results of shrink-wrap algorithm

Sales Person	Route generated	Distance to travel (units)
1	33-35-28-36-29-27-26-25-24-23-21-30-33	12.3227
2	33-48-34-12-11-10-9-5-1-4-2-3-33	13.3864
3	33-31-22-20-19-18-17-16-15-14-13-32-33	10.3390
4	33-47-46-7-49-6-59-58-50-51-57-45-33	12.5588
5	33-52-56-54-60-55-33	10.7616
6	33-44-43-53-42-40-41-39-37-38-33	10.1510

4.3 Shrink-Wrap algorithm

Shrink-Wrap algorithm is an algorithm to orient the nodes along a path. The nodes are mapped on polar coordinates, sorted by angle, then by distance (θ first, then ρ) and arranged in ascending order. This gives the path to be traversed within each cluster. The results are given in Table 2.

4.4 Simulated annealing for solving TSP

Simulated annealing (SA) is a generic probabilistic meta-algorithm for the global optimization problem, namely locating a good approximation to the global optimum of a given function in a large search space. The name and inspiration come from annealing in metallurgy, a technique involving heating and controlled cooling of a material to increase the size of its crystals and reduce their defects.

Simulated annealing can be aptly used for TSP. The output obtained using the cluster analysis heuristic is improved using route improvement heuristic which applied the meta-heuristic – Simulated annealing (SA).

The algorithm for SA is given below in Table 3 and the results obtained in this stage of the heuristic is given in Figure 3.

4.5 Tabu search for m TSP

Tabu search (TS) is a mathematical optimization method, belonging to the class of local search (LS) techniques. Tabu search enhances the performance of a local search method by using memory structures: once a potential solution has been determined, it is marked as "taboo" (thus the name, tabu search, since tabu and taboo mean the same thing) so that the algorithm does not visit that possibility repeatedly.

The basic principle of TS is to pursue LS whenever it encounters a local optimum by allowing non-improving moves; *cycling* back to previously visited solutions is prevented by the use of *memories*, called *tabu lists*, that record the recent history of the search, a key idea that can be linked to Artificial Intelligence concepts. The algorithm for tabu search is given in Table 4. The results of Tabu search is given in the Figure 4.

Table 3: Algorithm for Simulated Annealing

```

Algorithm1: Pseudocode for Simulated Annealing
X = Generate an Initial Feasible Solution;
C(X) = Compute initial cost of X;
best_cost = C(X);
T = Compute Initial Temperature;
While (stopping criterion not met)
    Repeat (pre-chosen number of times)
        Transition = Select a Transition from
            Neighbourhood (X);
        X' = Apply Transition(X, Transition);
        ΔC = Compute Change in Cost (X, X',
            Transition);
        p = generate random number (0,1);
        If ((ΔC < 0) OR (e-ΔC/T > p))
            X = X'
            C(X) = C(X) + ΔC
        End If;
        If (C(X) < best_cost) best_cost = C(X);
        End Repeat;
        T = Apply Cooling Function (T);
    End While;
Output best_cost;
End.

```

Table 4: Algorithm for Tabu search

```

Algorithm 2: Pseudocode for Tabu Search
X = Generate an Initial Feasible Solution;
C(X) = Compute initial cost of X;
best_cost = C(X);
Initialise tabu list T = ∅;
While (stopping criterion not met)
    For (s ∈ N(X))
        X' = Apply Transition(X, s, Transition);
        ΔC = Compute Change in Cost (X, X', s);
    End For;
    While (suitable neighbour not found)
        s ∈ N(X)
        If (s ∉ T)
            X = Apply Transition(X, s, Transition);
            T = T ∪ s;
            C(X) = Compute cost of X;
            found suitable neighbour = TRUE;
        Else
            If (aspiration(s) = TRUE)
                X = Apply Transition(X, s, Transition);
                C(X) = Compute cost of X;
                found suitable neighbour = TRUE;
            End If;
        End If;
    End While;
    If (C(X) < best_cost) best_cost = C(X);
End While;
Output best_cost;
End.

```

5 Implementation and Results

The problem was implemented using MATLAB 7.1 with Pentium IV processor system. After performing k -means clustering algorithm, Shrink wrap algorithm, simulated annealing and Tabu search were used. Table 5 shows the comparison of results for SA and TS.

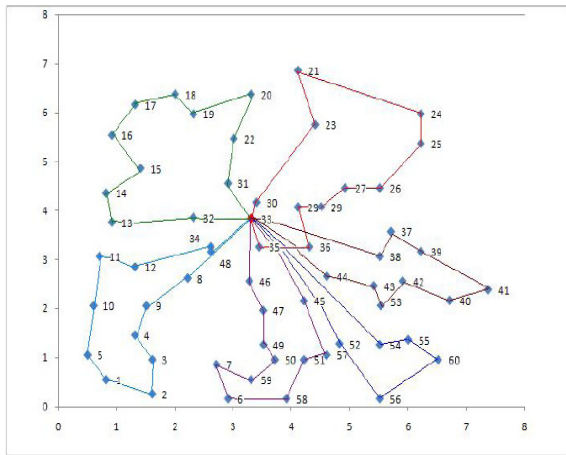


Figure 3: Sequence obtained from simulated annealing

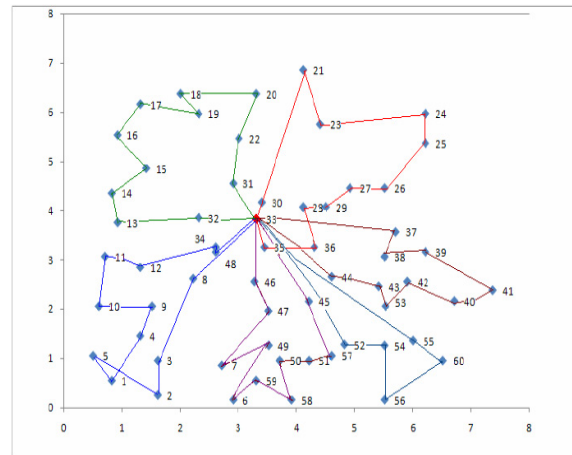


Figure 4: Sequence obtained from Tabu search

Table 5: Optimized Route Results After Applying SA and TS-Comparison of Distances

Salesman	Unoptimized Distance(units)		Optimized Distance(units)	
	Shrink wrap		Simulated Annealing	Tabu search
1	12.3227		11.2552	11.6938
2	13.3864		10.9832	12.7203
3	10.3390		10.3390	10.8722
4	12.5588		10.2939	11.5224
5	10.7616		10.1164	10.3594
6	10.1510		10.1510	10.2927
TOTAL	69.5195		63.1387	67.4608

6 Conclusion

From the results obtained, we find that k-means clustering proved to be effective as it was able to group the cities into clusters in an optimal manner and convergence took place in a short execution time and the optimal clusters were obtained. Between the sequences generated by both methods (SA and TS), the sequence generated by Simulated Annealing was more optimized. From the results it is evident that simulated annealing is better algorithm for a TSP than Tabu search.

By incorporating additional side constraints such as capacity, distance and time windows restrictions, it could easily be extended to a variety of vehicle routing problems (VRPs). Our future direction in this research would be, solving the same problem with multiple constraints like total time, cost and workload balance. The present problem was m TSP with single depot and this can be extended to m TSP with multiple depots. The work can be further extended by balancing the workloads of the salesmen by manipulating between clusters and reducing the standard deviation of the distance values. This will help in improving workers morale.

Acknowledgment

The authors would like to thank Associate editor and anonymous reviewers for their constructive suggestions and insightful comments that have helped to enrich the content and to improve the presentation quality of this paper.

References

- [1] A.E.Rizzoli, R.Montemanni, E.Lucibello, L.M Gambardella: Ant Colony Optimization for real world vehicle routing problems. *Swarm Intelligence, Springer New York*. 1(2): 135-151 (2007)

- [2] Allan Larsen: a study material, The Dynamic Vehicle Routing Problem. *IMM*. (2000)
- [3] Anderberg, M.R.: Cluster analysis for applications. *Academic press*, New York. (1973)
- [4] Arthur .E. Carter, Cliff. T. Ragsdale: A New Approach in solving Multiple Travelling Salesman Problem using Genetic Algorithm. *European journal of operational research*. 175(1): 246-257 (2005)
- [5] Bektas, T.: The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega*. 34(3): 209-219 (2006)
- [6] Weijun Xu, Xiaohong Cui: System Optimization on Distribution Center of Retail. *International Journal of Nonlinear Science*. 6(1): 79-85 (2008)
- [7] Bezalel; G. and Kizhanathan, S.: An optimal solution method for large-scale multiple traveling salesman problems. *Operations Research*. 34(5): 698-718 (1986)
- [8] C. B. Akki, S. M. Chadchan: The Survey of Handoff Issues in Wireless ATM Networks. *International Journal of Nonlinear Science*. 7(2): 189-200 (2009)
- [9] David S Johnson: Travelling Salesman Problem. *DIMACS Implementation Challenge, AT&T*. (2001)
- [10] Ding Chao, Cheng ye, He Miao: Two level genetic algorithm for clustered Travelling Salesman Problem with Application in large scale TSPs. *Tsinghua Science and technology*. 12(4): 459-465 (2007)
- [11] Gilbert K.C., Hofstra R.B.: A new multiperiod multiple traveling salesman problem with heuristic and application to a scheduling problem. *Decision Sciences*. 23: 250-259 (1992)
- [12] Hannes Schabauer, Erich Schikuta, Thomas Weishaupl: Solving Very Large Travelling Salesman Problems by SOM parallelization on cluster architecture. *In Proceedings of the Sixth international Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT)*. 954-958 (2005)
- [13] Zhongxin Liu, Xi Zhang, Zengqiang Chen, Zhuzhi Yuan: Exponential Stability Criteria for Feedback Controlled Complex Dynamical Networks with Time Delay. *International Journal of Nonlinear Science*. 7(1): 95-103 (2009)
- [14] Jian Huang, Jiazeng Wang, Zengrong Liu: Numerical Studies on the Epidemic Spreading on Correlated Networks. *International Journal of Nonlinear Science*. 5(1): 20-24 (2008)
- [15] Klaus Meer: Simulated Annealing Versus Metropolis for a TSP instance. *Information Processing Letters*. 104(6): 216-219 (2007)