A NEW SIMULATED ANNEALING APPROACH FOR TRAVELLING SALESMAN PROBLEM

Hüsamettin Bayram¹ and Ramazan Şahin²

¹Department of Industrial Engineering Hitit University, Çorum, Turkey
²Department of Industrial Engineering Gazi University, Ankara, Turkey
husamettinbayram@hitit.edu.tr, rsahin@gazi.edu.tr

Abstract- The aim of this study is to improve searching capability of simulated annealing (SA) heuristic through integration of two new neighborhood mechanisms. Due to its ease of formulation, difficulty to solve and various real life applications several Travelling Salesman Problems (TSP) were selected from the literature for the testing of the proposed methods. The proposed methods were also compared to conventional SA with swap neighborhood. The results have shown that the proposed techniques are more effective than conventional SA, both in terms of solution quality and time.

Key Words- Simulated Annealing, Travelling salesman problem, Roulette wheel selection, Meta-heuristics

1. INTRODUCTION

Meta-Heuristics are optimization techniques that start from an initial solution and search solution space iteratively improving the initial solution. Quality of solutions is improved during the search with regard to a given measure of quality. SA, Genetic Algorithms, Tabu Search are some of the popular nature inspired meta-heuristic algorithms which are used for the solution of combinatorial optimization problems. SA is based on the analogy between annealing of solids and solving of combinatorial optimization problems [1]. The SA algorithm is a stochastic meta-heuristic technique that uses randomized search and randomized acceptance methods which in return provides SA to escape from local minimum. Thus, SA enables effective searching of solution space using its specific mechanisms.

Thanks to the ease of formulation, difficulty to solve and various real life applications TSP is probably the most studied discrete optimization problem [2]. In TSP, number of cities (nodes) and the distances between them are known. The TSP is the problem of finding the shortest route that visits each city exactly once and returns to its origin. The TSP belongs to the class of NP-Hard problems. Therefore, for the large problems, heuristic approaches are the only viable solution techniques.

In this study, we introduce two modified versions of SA with new neighbor solution searching strategies. We tested these new approaches using several TSPs from the literature and we compared them to the conventional SA technique. Improvements will be discussed and the results of the experiments will be given in the following sections.
2. SIMULATED ANNEALING ALGORITHM

Mechanisms within SA prevent the search to quickly converge around a local minimum. During the search, SA not only accepts better solutions but also the worse solutions but with a decreasing probability. The probability of a worse solution to be accepted is determined by two parameters: the temperature and the difference between the objective function values (OFV) of current solution and neighbor solution. The aim of accepting worse solutions is to avoid converging of search to a local minimum. At higher temperatures, the probability of accepting worse solutions is much higher. But, as the temperature decreases, the probability of accepting worse solution decreases. In the following figure algorithmic steps of conventional SA is summarized.

Figure 1. Algorithm of Simulated Annealing
3. LITERATURE REVIEW

SA is a probabilistic meta-heuristic algorithm proposed by [3] and simultaneously by [4]. It has been proven that, SA is capable of solving many real life combinatorial optimization problems including scheduling problems [5], facility layout planning problems [6], assembly line balancing [7], vehicle routing [8] etc. As these studies are based on implementation of SA to a specific problem, some studies that contribute to the algorithmic steps of the SA are discussed in following paragraphs.

[9] proposed a fast SA approach where the cooling schedule of the SA is inversely linear in time. They showed their new cooling strategy is superior to the conventional SA technique. [10] suggested adaptive SA technique in order to reduce user defined search parameters. [11] studied Very Fast SA, where he introduced a new exponentially decreasing cooling schedule. [12] used threshold accepting method, which is principally simpler than conventional SA technique. They demonstrated their technique using TSP and showed that threshold accepting yields very near to optimum results for several known TSPs.

The authors of [13] presented a multi-objective Pareto SA approach, with the aim of finding set of efficient solutions formulti-objective combinatorial optimization problems.Objective weights are employed for the overall evaluation of solutions. [14] discussed parallel SA approach and they tested this technique using several TSPs from the literature. They have found out that the serial implementation of the SA is superior to conventional SA for the solving of TSP.

For the solution of TSP, the authors [15] improved adaptive SA with greedy search and they introduced three different mutation strategies for the generation of new solutions. Thus the convergence of SA is improved compared to several other algorithms in the literature. In [16] authors improved SA with greedy gradient mechanism and they implemented this technique for the solution of a course timetabling problem.

The above mentioned studies are a very little portion of studies that contribute to the SA algorithm. However, these techniques either do not use or partially use problem data in an intelligent and stochastic way for the generation of new solutions. The contribution of this study to the literature is twofold:

1. In order to generate good solutions, the proposed approaches intelligently use problem data or previous search experience by employing Roulette Wheel selection.

2. The proposed techniques do not sacrifice stochasticity and randomness while using problem data, which is essential for successful search of solution space.

4. NEW NEIGHBOUR GENERATION STRATEGY

As stated in previous sections, SA searches solution space by generating new solutions. One of the factors that impact choosing the strategy of neighbor generation is the representation scheme or in other words solution encoding. Representation scheme is the way of representing a candidate solution. In applications of SA to TSP, the most widely used and the simplest representation scheme is permutation encoding. In
permutation encoding the order of the numbers in the array represents the visiting order of the cities. For example, if the third element of the array is “5”, then it shows that the third city to be visited is the fifth city.

Selecting two random positions in permutation encoding representation scheme and swapping elements of these positions is the easiest and most widely used way of generating of neighbor solutions. In order to help to understand, in the Figure 2, two neighbor solutions (routes) for a TSP, the permutation encoding representation schemes that correspond to those solutions are given and calculations of their OFVs are described.

![Figure 2. Solution, solution representation and neighbor generation with swap operator](image)

As stated in previous paragraph, this neighbor searching structure is entirely random and does not take into account the lengths of entering arcs. We assume that the example given in the figure is for a symmetric TSP. In the figure, swap operator is applied to the positions 3 and 5. As it can be seen in the figure, after the swap neighbor generation operator (swap operator) edges 4-5, 5-1, 1-7 and 7-3 left the solution and edges 4-7, 7-1, 1-5 and 5-3 entered to the solution. Swap operator does not take into account the length of the entering edges. Therefore, a neighbor generation technique which intelligently uses the information of arch lengths without deteriorating diversification capability and randomness of SA is needed.

In Genetic Algorithms, Roulette Wheel selection or fitness proportionate selection is a widely used technique for the determination of parent solutions to be combined. As its name implies, in fitness proportionate selection, fitness values are assigned to the candidate solutions and a better fitness value increases the solution’s chance to be selected. Namely, every solution is associated with a probability which is proportional to its fitness value (OFV). Similarly, we propose to use Roulette Wheel as a selection mechanism of the edge that will enter to the solution. In the case of TSP, since the goal of TSP is to minimize the total length of route, shorter edges will enter to
the solution with higher probability. Introduction of edges to the solution is described in the following figure.

Let’s assume that after performing Roulette Wheel selection, 7-4 is selected as entering edge. Here, 7th node is origin (source), while the 4th node is destination (sink). Firstly, sink node (4th node) is separated from its preceding and succeeding nodes and connected to the source node (7th node) as the sink node. Namely, relative position of source node is saved and sink node is moved to the subsequent position of the source node.

5. PROPOSED METHODS

In this study, we propose two different extensions of SA. In the first method, entering nodes are selected from distance matrix, employing Roulette Wheel selection. In Roulette Wheel, shorter edges are selected with a higher probability. Selected matrices are introduced to the solution following the steps described in the last paragraph of the previous section. In brief, the first method is very similar to the conventional SA but the neighbor solutions will be generated with the help of Roulette Wheel selection.

The second method consists of two stages. In the first stage of this method conventional SA with swap operator is performed. During the first stage two different matrices are recorded. In the first matrix, the number of times each edge is part of a solution is recorded. In the second matrix, total OFV values for each of the edges for the solutions they enter during the search is recorded. At the beginning of the second stage of the second method, a third matrix is calculated for once. In this matrix, average OFV values of every edge are calculated using the first and the second matrices. The second matrix is divided by the first matrix, element by element. A lower average OFV value for an edge is an indicator of that edge can be a part of good solution with a higher probability. These average values are then used for the determination of entering edges in the second stage of the second method, instead of distance matrix. Similar to the first method, Roulette Wheel selection is applied in the second method using the third matrix and entering edges are determined and introduced to the solutions in every neighbor search. The methods are in summarized in below paragraph.

**Method 1:**

Step 1: Initialization

Step 1.1: Generate a random solution, Sol.
Step 1.2: Calculate solution’s OFV, \( OFV = f(Sol) \)
Step 1.3: Assign the initial “best solution” and its OFV value; \( Sol_{Best} = Sol; OFV_{Best} = f(Sol_{Best}) \).
Step 1.4: Set initial temperature; \( T = T_{max} \).
Step 2: Simulated Annealing
Step 2.1: Select a random edge employing Roulette Wheel selection and using the distance matrix.
Step 2.2: Generate a neighbor solution ($Sol_N$) by introducing new edge to the matrix and calculate neighbor solution’s OFV; $OFV_N = f(Sol_N)$.
Step 2.3: If $OFV_N < OFV$ accept new solution and move to the step 2.4. Else, generate a uniform random number between 0-1 ($u$) and calculate $\Delta = OFV - OFV_N$. If $u < e^{\Delta/T}$ accept new solution and move to the step 2.4. Else, reject solution and move to the step 2.5.
Step 2.4: Set $Sol = Sol_N$, $OFV = OFV_N$. If $OFV_N < OFV_{Best}$ set $Sol_{Best} = Sol_N$ and $OFV_{Best} = OFV_N$.
Step 2.5: If maximum number of searches reached for this temperature, then decrease the temperature; $T = T \times \lambda$, ($\lambda < 1$).
Step 2.6: If the minimum temperature is reached, stop the search and proceed to the Step 3. Else, go to the step 2.1.

Step 3: Report $Sol_{Best}$ and $OFV_{Best}$.

Method 2:
Step 1: Initialization
Step 1.1: Generate a random solution, $Sol$.
Step 1.2: Calculate solution’s $OFV$, $OFV = f(Sol)$
Step 1.3: Assign the initial “best solution” and its OFV value; $Sol_{Best} = Sol$; $OFV_{Best} = f(Sol_{Best})$;
Step 1.4: Set initial temperature; $T = T_{max}$.
Step 2: Simulated Annealing
Step 2.1: Select two random positions in the current solution ($Sol$).
Step 2.2: Generate a neighbor solution ($Sol_N$) by swapping the elements of these positions calculate neighbor solution’s $OFV$; $OFV_N = f(Sol_N)$.
Step 2.3: If $OFV_N < OFV$ accept new solution and move to the Step 2.4. Else, generate a uniform random number between 0-1 ($u$) and calculate $\Delta; \Delta = OFV - OFV_N$. If $u < e^{\Delta/T}$ accept new solution and move to the step 2.4. Else, reject solution and move to the step 2.5.
Step 2.4: Set $Sol = Sol_N$, $OFV = OFV_N$. Increase edge count matrix’s elements (matrix 1), that correspond to the edges in $S_N$, by 1. Namely, if there is an edge $i$-$j$ in $Sol_N$ increase the element in the $i^{th}$ row and in $j^{th}$ column by 1. In addition, add $OFV_N$ to the same elements of the matrix 2. If $OFV_N < OFV_{Best}$ set $Sol_{Best} = Sol_N$ and $OFV_{Best} = OFV_N$.
Step 2.5: If maximum number of searches reached for this temperature, then decrease the temperature; $T = T \times \lambda$, ($\lambda < 1$). Else go to step 2.1.
Step 2.6: If the minimum temperature is reached, stop the search and proceed to the Step 3. Else, go to step 2.1.

Step 3: Divide matrix 2’s every element by same element of the matrix 1 and create a new matrix that stores an average value for every edge (Matrix 3).
Step 4: Initialization
Step 4.1: Generate a random solution, $Sol$.
Step 4.2: Calculate solution’s $OFV$, $OFV = f(Sol)$
Step 4.3: Assign the initial “best solution” and its OFV value; $Sol_{Best} = Sol$; $OFV_{Best} = f(Sol_{Best})$.

Step 4.4: Set initial temperature; $T = T_{max}$.

Step 5: Simulated Annealing

Step 5.1: Select a random edge employing Roulette Wheel selection and using the matrix $3$.

Step 5.2: Generate a neighbor solution ($Sol_N$) introducing new edge to the matrix and calculate its length; $OFV_N = f(Sol_N)$.

Step 5.3: If $OFV_N < OFV$ accept new solution and move to the step 5.4. Else, generate a uniform random number between 0-1 ($u$) and calculate $\Delta = OFV - OFV_N$. If $u < e^{\Delta/T}$ accept new solution and move to the step 5.4. Else, reject solution and move to the Step 5.5.

Step 5.4: Set $Sol = Sol_N$, $OFV = OFV_N$. If $OFV_N < OFV_{Best}$ set $Sol_{Best} = Sol_N$ and $OFV_{Best} = OFV_N$.

Step 5.5: If maximum number of searches reached for this temperature, then decrease the temperature; $T = T \lambda$, ($\lambda < 1$).

Step 5.6: If the minimum temperature is reached, stop the search and proceed to the step 6. Else go to the step 5.1.

Step 6: Report $Sol_{Best}$ and $OFV_{Best}$.

![Figure 3. Graphical illustration of the proposed methods](image)

6. EXPERIMENTS

For the testing of these two proposed methods, 11 example problems from the literature were used. Then these two methods are compared to conventional SA with swap operator. Experiments were conducted in two stages. In the first stage of the experiments three methods were tested without time limit and with the same SA parameters. In the second stage, time limited versions of the techniques were run and
compared. All these methods were written in MATLAB programming language and every method was run 30 times for every problem. For the comparison of the techniques, relative error of the average results with respect to the optimal solution was used. Calculation of average error rate is described below: Results of the first stage of the experiments are given in Table 1.

\[
\text{Average Error Rate (\%)} = \frac{(\text{average OFV}) - (\text{OFV optimal})}{(\text{OFV optimal})} \times 100
\]

Table 1. First stage of the experiments

<table>
<thead>
<tr>
<th>Method</th>
<th>Cooling rate (alpha)</th>
<th>Min</th>
<th>Max</th>
<th>Avg</th>
<th>Std. dev.</th>
<th># of Solutions searched</th>
<th>Time (avg.)</th>
<th>Avg. Err. (%)</th>
<th>Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gr17 [17]</td>
<td>1000000</td>
<td>20</td>
<td>0.99</td>
<td>2083</td>
<td>2149</td>
<td>2092.5</td>
<td>14.73</td>
<td>0.69</td>
<td>0.36</td>
</tr>
<tr>
<td>Gr21 [17]</td>
<td>100000</td>
<td>20</td>
<td>0.99</td>
<td>2707</td>
<td>3000</td>
<td>2738.6</td>
<td>72.69</td>
<td>0.75</td>
<td>1.17</td>
</tr>
<tr>
<td>Gr24 [17]</td>
<td>100000</td>
<td>20</td>
<td>0.99</td>
<td>1272</td>
<td>1408</td>
<td>1322.7</td>
<td>37.65</td>
<td>0.75</td>
<td>3.99</td>
</tr>
<tr>
<td>Bayg29 [18]</td>
<td>100000</td>
<td>20</td>
<td>0.99</td>
<td>1610</td>
<td>1813</td>
<td>1658.2</td>
<td>53.50</td>
<td>0.80</td>
<td>2.99</td>
</tr>
<tr>
<td>Dantzig42 [19]</td>
<td>100000</td>
<td>20</td>
<td>0.99</td>
<td>699</td>
<td>740</td>
<td>713.5</td>
<td>12.99</td>
<td>9.69</td>
<td>2.08</td>
</tr>
<tr>
<td>KroA150</td>
<td>200000</td>
<td>20</td>
<td>0.99</td>
<td>7542</td>
<td>8719</td>
<td>8049.2</td>
<td>281.24</td>
<td>230140</td>
<td>11.36</td>
</tr>
<tr>
<td>KroA50</td>
<td>200000</td>
<td>20</td>
<td>0.99</td>
<td>767</td>
<td>797</td>
<td>727.9</td>
<td>26.29</td>
<td>230140</td>
<td>13.69</td>
</tr>
</tbody>
</table>

As it can be seen in Table 1, in the first stage of the experiments the results obtained using first and second methods are better compared to the results reached using conventional SA with swap operator. However, first and second methods’ running time was up to 18 times longer than that of conventional SA due to the computational complexity. In the second stage of the experiments, conventional SA’s epoch length is increased in such a way that its running time slightly exceeds that of first and second methods. The aim of that is to compare proposed methods’ effectiveness in a limited time. For this reason, in the second stage of experiments only conventional SA was run with increased epoch length. The results of the second stage of the experiment are given in Table 2.
In the second stage of the experiments, conventional SA with swap operator was performed. In order to increase its running time and improve the quality of solutions, epoch lengths were increased. As it can be seen in Table 1 and Table 2, results obtained using the proposed methods are superior to the results obtained using conventional SA with swap operator. It can be inferred that, neighbor generation through intelligent introduction of edges increases the capability of SA to reach better solution. In addition, second method was also slightly better than the first method. But, it must be kept in mind that, the times given for second method are not the actual running times of second method. The time required for the generation of matrix 1 and matrix 2 are excluded. That means, it cannot be asserted that second method is better than first method.

### 7. CONCLUSION

This paper introduces two new solution approaches for TSP based on SA. These techniques employ Roulette Wheel selection strategy for the selection of entering edges to a solution representation. The first method uses distance matrix and the second method uses a new matrix which is created using the average OFVs of previously visited solutions. Then, the proposed methods were tested on 11 well known benchmark problem sets from literature. After the experiments, it is found out that the proposed methods are superior to conventional SA with swap parameter in reaching optimal and suboptimal solutions. The main advantage of these techniques is improved convergence capability by using problem data and without losing inherent stochasticity in the SA technique. The paper has also revealed that the Roulette wheel selection of genetic algorithm is also applicable in SA with success.

### 8. REFERENCES


