Simulated Annealing for Traveling Salesman Problem with Hotel Selection for a Distribution Company Based in Mexico

Raúl Jiménez-Gutiérrez*, Diana Sánchez-Partida, José-Luis Martínez-Flores, Eduardo-Arturo Garzón-Garnica

Department of Logistics and Supply Chain Management, UPAEP University, 17 Sur 901, Barrio de Santiago, CP 72410 Puebla, Puebla, México

**Abstract**

A distribution company in Mexico covers the travel expenses for 21 sales representatives. Currently, the routes they follow are not established clearly, which can lead to high costs in this subject. A reduction of such cost is sought after, by optimizing the routes for each one of them. The following research finds an improvement on the routes for the sales representative of a distribution company in Mexico. It was done by using the Traveling Salesman Problem with Hotel Selection or Base selections via a Simulated Annealing algorithm. The results show an improvement in a reasonable timeframe by using the Simulated Annealing. It also shows that the maximum process time was of 156.63 minutes, and the least amount of improvement was 24.44% over the current route selection. Applying this model will be beneficial for the company as the company is trying to reduce costs related to the sales representatives such as; fuel cost, hotel cost, and travel expenses.

**Keywords:**

- TSP
- TSPHS
- TSP-HOBS
- Simulated Annealing
- Metaheuristic

1. Introduction

For the agro-industry, to satisfy the specific demands of customers using different strategies in product and service has proven successful in an increasingly competitive global market.

Logistics has been an essential part of the success of the companies. Nowadays, there is more focus on how companies can optimize their resources and do more with less, a perfect balance between efficiency and customer service. According to [1] logistics are vital because it creates value, for customers and suppliers of the firm, and value for the firm’s stakeholders. Value in logistics is expressed in terms of time and place.

Logistics have several improvement methods that could be implemented, not only in the logistics areas of the companies but also in commercial areas. Mexico is not an exception when TSP models can be helpful as the lack of road infrastructure and the long distances, besides the facts mentioned there some companies trying to make savings by reducing the sales force and extending the region coverage of the sales representatives.

As an example, we can find the Traveling Salesman Problem (TSP), where rapid growth of the companies raises the need to implement commercial strategies to visit their clients with the highest possible efficiency and quality but with the least cost. It is at this point is where mathematical models show up to create a clear vision and achieve such goals. Moreover, in the case of the TSP, it can help the companies to reduce the fixed costs for the sales representative such as:

1. Fuel costs.
2. Pay toll costs.
3. Maintenance costs.
4. Accommodation costs.
5. Meal expenses.
6. GPS costs.
7. Other travel expenses.

To develop study cases related to having route efficiency is essential for the companies established in Mexico. According to the National Association of supermarkets and department stores, during 2019, Mexico has grown 7.7% [2]. Furthermore, part of this growth is due to the effort of the sales force.

Therefore, for the current research, a variant for the TSP will be implemented, the Traveling Salesman Problem with Hotel Selection (TSPHS) with Simulated Annealing. The main objective for a TSPHS is to have an efficient route, including the stops where the sales representative will rest after their work shifts as they need, in many cases, to visit several clients.
2. Problem Description

A distribution company based in Mexico is responsible for delivering products to a nationwide authorized dealer’s network. Goods are sent using established couriers, and the distribution center is based in the central region of the country. All sales are made through authorized dealers to assure the quality of service. The company tries to keep a high level of service for the dealers and has several sales representatives throughout the country. The role of each sales representative is to establish a sales plan with each dealer, take orders, assure such orders have been fulfilled, and present the dealers with technical information and training, both for existing products as well as new products. The company decided to divide the country’s sales territory into eleven commercial zones to have an efficient approach, and to know and cover the needs of the dealer’s network. Fig. 1 shows the current zones within the country, the distribution company provided the information as part of its current structure of commercial zones.

![Figure 1: Sales Representative Zones](image)

Each commercial zone has a sales representative in charge. The sales representative, the primary task is to visit each dealer of the zone he is in charge of and know the needs of each dealer. He must provide feedback about the sales plan that the dealer has for the year, as well as other tasks that include a demonstration of new products, special deliveries (this term is used when a sales representative brings support to the final user and explains how to assemble, use and give maintenance) and training.

It is worth to mention that the company covers the travel expenses for the route or the visits that they need to do. It means that the company covers the gasoline, toll payment, hotels, meals, and other possible expenses that the sales representative incurs during the visits in its entirety. The sales representative needs to make all of the visits as traveling during his working hours, which are set to nine working hours per day. The information provided by the company did not include the time each sales representative spends with the dealer/client. However, the variation in travelling time between the calculated average and the real one was found to be enough to cover the time of the visit. It was decided to leave the visit time out of the calculations, to ease the solving of the model.

Despite the facts above, there is no specific route for the sales representative. Therefore, the order in which to drive to each dealer store to make their visits is entirely up to the sales representative. In several cases, they pass through the same point twice. Being inefficient on the route design could increase the expenses for the company. As an example, the sales representative in charge of the south region is in charge of 19 dealers. Back in 2019, he drove 57081 km, equivalent to MXN 98,011.00 in gasoline expenses only.

The main objective of this research is to find an improvement in the efficiency of the routes that each sales representative travel among their specific commercial zones, using a well-known combinatorial optimization problem, the Traveling Salesman Problem with hotel selection (TSPHS) solved by a simulated annealing method, focused on finding the following objectives:

- Improvement of the routes.
- Find the hotel that allows the sales representative to rest or, in some cases, go back to their homes but still being efficient on the route.

As part of the improvement strategy to have better routes, the distribution company conducted and applied a survey to the sales representatives to know their preferences about staying in a hotel or at their home. All sales representatives agreed that they prefer to get back to their homes, but they know that, in most cases is not possible to return to their homes due to the long distances that they need to drive.

<table>
<thead>
<tr>
<th>ZONE</th>
<th>HOTEL</th>
<th>HOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>M11</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>M12</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>M13</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>M14</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>M21</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>M22</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>M23</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>M24</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>M31</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>M32</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>M33</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

The sales representatives commented as well that, in some cases, they suffer from stressful situations, as there are several times when they do not have the chance to get back to their homes, and when they try to book a hotel, there are no available rooms. According to [3], sales representatives are susceptible to stress by the nature of their job. Therefore, this research includes the hotel and the home of the sales representative that for this purpose is called base.
The results of the aforementioned internal survey were compiled into a table with the number of hotels commonly used in each zone. The table is the following:

It is worth to mention that, the agent has a list of authorized hotels that the company selected previously. All of those hotels are in the same price range; therefore, the sales representative can choose either one, without it representing a considerable difference in cost. Hence the limitations for the research will be to establish the routes for each sales representative.

3. Literature Review: Travelling Salesman Problem with Hotel Selection and Simulated Annealing.

The problem of this research has been studied in several research cases, and it is well described in the literature. The TSP is one of the most well-known combinatorial optimization problems, where the main idea is to find an optimal solution to avoid passing the same point twice. Over the years, several methods based on deterministic or probabilistic heuristics have been proposed to solve the known problem. The method includes branch-and-bound, particle swarm optimization, neural network, and simulated annealing [4].

In simple words, the TSP consists of determining a minimum distance circuit passing through each vertex once and only once. Such a circuit is known as a tour or Hamiltonian circuit [5].

A formal definition of TSP, according to [6]:

Let $K_n = (V_n, E_n)$ be the complete undirected graph with $n = |V_n|$ nodes and $m = |E_n| = \binom{n}{2}$ edges. An edge $e$ with endpoints $i$ and $j$ is also denoted by $ij$, or by $(i, j)$. We denote by $\mathbb{R}^{E_n}$ the space of real vectors whose components are indexed by the elements of $E_n$. The component of any vector $z \in \mathbb{R}^{E_n}$ indexed by the edge $e = ij$ is denoted by $z_e, z_{ij}$, or $z(i, j)$.

Now, considering that the salesperson, in some cases, needs to travel long distances, the need for a hotel stay arises. Such stay is an everyday activity, to end his daily route and start his route the next day to visit another dealer. Selecting a hotel to end/rest/start the route converts the research into a traveling salesman problem with hotel selection.

According to this research, there are several cases where the Traveling Salesman Problem with Hotel Selection (TSPHS) has been studied; Vansteenwegen [7] introduced the concept of TSPHS, the difference between TSP and TSPH, the last one is where a mobilized entity (salesman) cannot visit all locations (customers) at once without taking any break (break at a hotel location). It is because the salesman has limited working hours; the target at the end is to reduce the total travel time.

The TSPHS can be defined [7] on a complete graph $G = (V, A)$ where $V$ is the set of vertices and $A = \{(k, l) \mid k, l \in V, k \neq l\}$ is the set of arcs. There are $s$ available hotels (which are indexed through $k = 1 \ldots s$) and $n$ customers (which are indexed through $k = (s + 1) \ldots (S + n)$) in the set $V$. Each customer $k$ is assigned service or visiting time $T_k$ where $T_k$ is $k = 0, 1 \ldots, \ldots$. The time $kt_c$ needed to travel between locations $k$ and $l$ is known for all $(k, l) \in A$.

Several examples of TSPHS can be found in the literature as well as new heuristics solution methods. As it is mentioned in [8] that the TSPHS is a difficult optimization problem that arises in many practical situations, in this case, the authors gather the data into four sets:

<table>
<thead>
<tr>
<th>Set</th>
<th>Number of Customers</th>
<th>Number of hotels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48-288</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>52 - 1002</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>52 - 1002</td>
<td>10</td>
</tr>
</tbody>
</table>

As can be seen, it is similar to the distribution company based in Mexico, but instead of having four sets, it has eleven commercial zones with a specific number of customers and hotels.

The new metaheuristic that they implemented is a two-phase multi-start procedure that combines a simple GRASP (Greedy randomized adaptive search procedures) and a variable neighborhood descent to improve the solution. They find that for smaller instances, the method was able to find either optimal solutions or best solutions, but also that the difficulty increases if the number of hotels grows.

Another example of solving the TSPHS with a new metaheuristic solution can be found with a Migrating Bird Algorithm (MBO) [9], and the solution approach is based on the flight formation of migrating birds, using this method it is shown that it is efficient for solving the problem in a reasonable computational time. Furthermore, in this paper, it is mentioned that simulated annealing was run, and it got results in a reasonable timeframe as well. For the TSPHS with Migration Bird Algorithm, the result shows that the gap between the MBO and the exact solution is acceptable.

Other methods available include Particle Swarm Optimization [10], [11], Artificial Bee Colony [12], Ant Colony [13], Crow Search Optimization [14], Cat Swarm Optimization [15], Grey Wolf Optimization [16], Whale Optimization [17]. The advantages/disadvantages of these methods are all well documented in literature. Some of these methods have successfully been used to solve TSP [12], [13], [16]. And even some have been applied in combination with Simulated Annealing [11]. In previous (unpublished) work by the authors, the Simulated Annealing algorithm was programmed and tested. The advantages and disadvantages of applying each of the aforementioned methods were evaluated. But the largest consideration was still the cost of implementing, or programming, such algorithms. In all the cases, the cost, expressed mainly in programming time, of implementing the newer algorithms was higher than:

www.astesj.com

502
1. The available time the authors could allocate for the research.
2. The deadline the company had to respect.

Both these factors resulted in a cost that was higher than the benefits of implementing a newer algorithm, even if it would eventually yield better results. Still, an improvement could be obtained by implementing a simpler, cheaper (in this case) to apply method.

A simple example of solving the TSPHS [18] that was formulated as a mixed-integer linear programming (MILP).

\[
\begin{align*}
\min & \quad \sum_{d=1}^{m} \sum_{k=0}^{s+n} \sum_{l=0}^{s+n} x_{k,l,d} C_{k,l} \\
\text{s.t.} & \quad \sum_{i=0}^{s} x_{0,l,1} = \sum_{k=0}^{s+n} x_{k,0,m} = 1 \\
& \quad \sum_{h=0}^{s+n} \sum_{k=0}^{s+n} x_{k,h,d} = \sum_{h=0}^{s+n} \sum_{l=0}^{s+n} x_{h,l,l} = 1, \quad d = 1, \ldots, m \\
& \quad \sum_{i=0}^{s+n} x_{k,i,d} = 1, \quad i = (s+1), \ldots, s+n \\
& \quad \sum_{k=0}^{s+n} \sum_{l=0}^{s+n} x_{k,l,d} = s+n, \quad d = 1, \ldots, m; i = s+1, \ldots, s+n \\
& \quad \sum_{k=0}^{s+n} \sum_{l=0}^{s+n} x_{k,l,d} (C_{k,l} + T_{d}) \leq C, \quad d = 1, \ldots, m \\
& \quad u_{i} - u_{j} + 1 \leq n - 1 \left(1 - \sum_{d=1}^{m} x_{i,j,d}\right) \quad i = s+1, \ldots, s+n; j = s+1, \ldots, s+n \\
& \quad x_{i,j,d} \in \{0,1\}, \quad d = 1, \ldots, m; i = 0, \ldots, s+n; j = 0, \ldots, s+n \\
& \quad u_{i} \in 1, \ldots, n, \quad i = s+1, \ldots, s+n
\end{align*}
\]

For this instance, [18] the goal of the TSPHS is to determine a tour of minimal length that visits all n customers. The tour is composed of connected trips with limited length, and every trip should start and end in one of the available hotels. The hotels that are selected determine to a great extent, the length of the total tour.

Combining the TSPHS with another strategy can yield an advantage, in recent research, it can be found that the simulated annealing as a complementary for effective results. In [19] can be found that Kirkpatrick introduced the implementation of simulated annealing. The process begins by considering a solution space \(S\) of an exclusive tour through the set of given cities or points \(1, 2, \ldots, n\), with a set of updated solutions \(x_{t}^{i}\) created by randomly switching the orders of two cities. The energy function, or fitness function, which represents the length of the route \(x_{t}\), is denoted by \(f(x_{t})\). The relative change \(\Delta \in \) in cost \(f\) between \(x_{t}\) and \(x_{t}^{i}\) is expressed as 

\[
\Delta f = \frac{f(x_{t}^{i}) - f(x_{t})}{f(x_{t})}.
\]

Besides, it is mentioned in [20] that using the potential usefulness of a fixed temperature algorithm is considered suitable for small problem instances. On the same document [20], where they implement a Simulated Annealing (SA) to solve a TSP, the optimal fixed temperature is chosen experimentally by running simulated annealing many times at each of several fixed temperatures and determining which temperature is best, according to an appropriate optimality criterion. As a result, they see that fixed-temperature simulated annealing tends to outperform the use of a fast cooling schedule, but for TSP instances of size less than 150 cities [20].

4. Solution Approach

The only information the authors had in hand was the list of coordinates of the clients, hotels, and base of each zone. However, a detailed record of the agents’ travels was not available. As the travel time was to be limited, because of the maximum work time, and to be able to know at which moment a hotel had to be selected, some kind of transformation was needed. The geographical data had to be converted to travel time. The process is explained next.

A sample set of location pairs, covering as many zones as possible, was selected. For each pair of locations, the traveling time was sought, using Google Maps®. Also, the Euclidean distance for that pair was computed. For all pairs, the relationship between the Euclidean distance, the Google Maps® distance, and the travel time was set up in a table. The traveling time varied between 3 minutes and 917 minutes. The Google Earth distance varied between 650 meters and 1,153,000 meters. The Euclidean distance varied between 413 meters and 954,496 meters. The proportion varied between 0.0006 and 0.007 and averaged 0.0013. Therefore, multiplying the Euclidean distance by the proportion’s average should yield an approximation to the travel time between both locations.

A simulated annealing method was implemented instead of getting an exact solution due to the fact - that previous research shows that simple TSP models with 50 locations take too much time to get a solution or, even worse, do not get a solution using currently available computing equipment. Additional to this, using metaheuristics yields a result reasonably close to those of the exact methods [21]. As the solution requires the addition of

www.astesj.com
the hotel selection, which adds even more complexity to the model solving, a quicker solving method was needed. Such is the reason for the use of heuristics instead of trying to obtain an exact solution.

As was already mentioned, a Simulated Annealing algorithm was used. It was programmed in GNU Octave®. The objective of such an algorithm was to find an improvement in the total time of travel for each sales agent. It was, therefore, improving the operating costs of the company. Some considerations had to be taken to achieve such a goal.

If the agent has to choose a hotel and the travel time from his current location to the nearest hotel is larger than the time to his base, including a 30-minute tolerance, the agent will return to his base.

The sum of all the distances traveled, including the times the traveler returns to his base, were considered to evaluate the objective function. According to [22], SA simulates a collection of atoms in equilibrium at a set temperature. SA starts with an initial state. A random change is made to this state, and the change in energy, ΔE, is calculated. If the new state has lower energy than the previous state, i.e.

ΔE ≤ 0

the new state is carried into the next iteration. However, if the new state has high energy, it is accepted with probability P(ΔE) = exp(−ΔE / K_B T) where T is the temperature and K_B it is Boltzmann’s constant. This acceptance of an uphill move (a move with ΔE ≥ 0) is analogous to a higher energy molecule knocking lose a molecule trapped in a state of excess energy Determining the cooling (annealing) schedule in step 3(f) of SA is crucial to the success of the algorithm. The chance of accepting an energy-increasing move is inversely proportional to ΔE and proportional to T, which decreases with time.[22]

The following steps were taken into account in order to generate a new solution,

• The current best solution is obtained, and the hotels and returns to the base are removed. The remaining list includes only client locations.
• Two points are obtained randomly, and an exchange, or flip, is performed.
• The base is added at the beginning and end of the route.
• Each leg is traveled, calculating the travel time through the previously obtained time-distance relationship.
• When the accumulated time to arrive at the next client exceeds the maximum journal duration, a hotel or base is added. It is done by obtaining the middle point between both clients and finding the hotel or base closest to such point. If the travel time to the base is lesser than to a hotel, or even if it is more extensive, up to a tolerance of 30 minutes, the base is selected. Then the resting point is added, and the journal time reset.

5. Problem Solution

The first solution was considered as the order in which each client appears in the database. The SA model was run for each zone, finding an improvement in each case. A graph showing the behavior of the model in one of such runs is shown in Figure 2. Improvements were found for all the zones, the execution time and improvement of all the zones can be observed in Table 3.

![Figure 2: Model Behavior](image)

<table>
<thead>
<tr>
<th>Zone</th>
<th>Init OF</th>
<th>Final OF</th>
<th>Improvement</th>
<th>Runtime (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M11</td>
<td>54.3782</td>
<td>38.4468</td>
<td>29.29%</td>
<td>114.77</td>
</tr>
<tr>
<td>M12</td>
<td>76.8986</td>
<td>56.0846</td>
<td>27.06%</td>
<td>113.06</td>
</tr>
<tr>
<td>M13</td>
<td>34.2869</td>
<td>19.4949</td>
<td>43.14%</td>
<td>63.25</td>
</tr>
<tr>
<td>M14</td>
<td>102.632</td>
<td>73.683</td>
<td>28.20%</td>
<td>231.82</td>
</tr>
<tr>
<td>M21</td>
<td>39.8978</td>
<td>18.2448</td>
<td>54.27%</td>
<td>69.53</td>
</tr>
<tr>
<td>M22</td>
<td>41.5431</td>
<td>22.5259</td>
<td>45.77%</td>
<td>89.24</td>
</tr>
<tr>
<td>M23</td>
<td>18.4146</td>
<td>12.1156</td>
<td>34.20%</td>
<td>77.32</td>
</tr>
<tr>
<td>M24</td>
<td>130.139</td>
<td>84.894</td>
<td>34.76%</td>
<td>266.01</td>
</tr>
<tr>
<td>M31</td>
<td>54.583</td>
<td>41.2396</td>
<td>24.44%</td>
<td>156.63</td>
</tr>
<tr>
<td>M32</td>
<td>28.2553</td>
<td>14.4192</td>
<td>48.96%</td>
<td>68.57</td>
</tr>
<tr>
<td>M33</td>
<td>75.500</td>
<td>48.4674</td>
<td>35.80%</td>
<td>177.82</td>
</tr>
</tbody>
</table>

An example of the resulting routes is shown in Figure 3, where the return of the agent to his base can be appreciated. To ease the reading of this document and to avoid disclosing information that could lead to the identifying of the company or disclosing its clients, the other tables and routes will not be shown.

![Figure 3: Sample Route](image)

The model was run on a laptop PC. Intel® Core i5, 8Gb RAM. As shown in table 3, the execution times varied from 63.25
minutes to 266.01 minutes, or from roughly one hour to four and a half hours. The full execution of the model for all zones did not take more than two full days. An attempt to solve the model precisely by the use of a linear solver was also made, but the execution took more than 100 hours without results, so the SA results were the only ones available.

6. Conclusions

An improvement in the travel time and hotel stays of the sales representatives for the distribution problem of an international company with a presence in Mexico was sought. The scope of the problem included all the Mexican territory, divided into eleven commercial zones. A Simulated Annealing method was selected to try and find better results than the current ones in a reasonable time frame.

The only information available was the location of each client, hotel, and base, but the problem needed a way to relate those to travel times. A relationship was obtained by calculating the travel time within a small number of locations. In this way, the relationship between Euclidean Distance and Travel Time was obtained.

A Travelling Salesman Problem with Hotel or Base Selection was implemented, using the aforementioned Simulated Annealing method, for each of the commercial zones. The model was run, and new routes were obtained for each zone. The maximum process time was of 156.63 minutes, and the least amount of improvement shows a 24.44% improvement over the current route selection.

6.1. Scope for future research

Using more precise mathematic models to find the relationship between the distance and the travel time can improve the efficiency of the calculation of travel time when only the coordinates are known. It could mean that more straightforward calculations would yield more precise results when travel time is needed.

References