Evolutionary and Metaheuristic Approach for the Traveling Salesman Problem: A Comparative Study

Akimuddin Aslam Shaikh
School of Computing
National College of Ireland
x22123245@student.ncirl.ie

***Abstract*—The Traveling Salesman Problem (TSP) is a well-known and a historical combinatorial optimization problem. It aims to find the shortest possible paths that a salesman can take as he travels from one point to another and then coming back to the starting point. The practical applications of problem are planning, logistics, and various fields in computer science. This report covers the historical progress, implements 2 metaheuristic algorithms, and compares the performance of algorithms like Integer Programming, Genetic Algorithm, and Simulated Annealing contributing to TSP solutions. These algorithms were then run on datasets of 29, 40, and 80 cities of Bavaria to find the shortest distance round trip and also consider execution time. The report made an attempt to improve upon the results by enhancing Genetic Algorithms and Simulated annealing. However, results showed that Integer Programming yields optimal solutions consistently. While, Genetic Algorithms and Simulated Annealing seem to strike a balance between accuracy and efficiency.**

Keywords: Traveling Salesman Problem, Integer Programming, Genetic Algorithm, Simulated Annealing, Historical review

I. Introduction

The Traveling Salesman Problem (TSP) is a long-standing, well-studied problem in the combinatorial optimization domain. It pose as a simple yet difficult task due to its NP-hard nature, finding the shortest path for a salesman to visit each city exactly once, and return to the starting point. It has served as a benchmark for optimization techniques.

This study examines the roots and significance of TSP which dates back to a manual published in 1832 [1] listing difficulties encountered when planning routes that were formulated mathematically during the 19th century. The project studies early researches by the noble academicians who have contributed significantly to this problem.

The project then proceeds to solve TSP using two different modern meta-heuristic algorithms. It implements and compare the performance of three metaheuristic algorithms performed: Integer Programming, Genetic Algorithm, and Simulated Annealing. It also made an attempt to make the results of the Genetic Algorithm and Simulated Annealing better by creating an upgraded version of each. The basic Genetic Algorithm used common selection, crossover, and mutation operators [2]. The advanced one, added an edge recombination crossover and changed mutation rates to get better outcomes. The basic Simulated Annealing algorithm used a simple cooling schedule, whereas, the advanced version made this better with the 2-opt local search method and changing perturbation strengths to get out of local minima more. Each of the algorithms was tested on datasets of 29, 40, and 80 Bavarian cities. The performance evaluation was based on the length of the shortest roundtrip found and the computational time required.

The report is structured as follows:
1. The project studied the history and advancement in the researches of the implemented algorithms.
2. Implementation of methods and structure in chosen algorithms inspired by the literature, and advance version of Genetic algorithm and Simulated Annealing.
3. Finally, the report analyzed and discuss the performance of developed algorithms, the future scope of the work, and benefits and applications of our work to the real world for the future researchers.

This structured approach provides an in-depth information about algorithms and provide valuable key points in TSP solutions development that align both with historical perspectives and practical applications.

II. History of the Problem

The origin of Optimization problem was known to be appeared in 1832 [1] from a manual book titled Der Handlungsreisende — wie er sein soll und was er zu thun hat, um Aufträge zu erhalten und eines glücklichen Erfolgs in seinen Geschäften gewiß zu sein, written by an "old Commis-Voyageur." It describes the the problem of optimizing travel routes, visiting multiple locations in one go, and returning to the starting point in the shortest distance possible. The book suggests five routes for the traveling salesman through Germany and Switzerland, one of which approximates a modern shortest path solution closely. However, This manual lack mathematical formulation and just highlight the practical importance of optimization.

A. Major Publication

The paper [3] published in 1954 addressed the TSP, involving 49 cities, one from each state in US and Washington, D.C. Three methodologies were implemented namely Linear Programming, Cutting-Plane method and Graphical interpretation. Cutting-plane methods involve adding hyperplanes that is iteratively adding constraint to linear programming to eliminate non-optimal solutions and converge on the optimal tour. Graphical interpretation was employed as they reduce the problem dimensionality by applying a triangular array of distances and mapping n-space onto a lower-dimensional space. The authors successfully solve the 49-city TSP problem, proving that their approach can yield the optimal solution. While the authors also acknowledged that they did not solve all the aspects of TSP, However, their method still have a profound effect on combinatorial optimization and integer programming.

1) History of Genetic Algorithm: The research by [2] was a serious contribution specifically to the application of GAs for finding solutions to TSP. It showed that GAs are capable of finding a variety of near-optimal solutions to very complex combinatorial problems. It does all this through a framework in methodology, precisely detailing representation—chromosome as tours, selection, crossover, and mutation—made with regard to the TSP problem. The research include practical results that show their genetic algorithm's performance with appropriate operators and parameter tuning on several instances of TSP. The results obviously attest to the efficiency and effectiveness of the approach compared to traditional methods of optimization. This work would, in the future, act as a seedbed for more research in the application of hybrid genetic algorithms in other optimization problems, hence influencing subsequent works in many other fields. The paper enriches the interdisciplinary nature by integrating concepts from biology and computer science, advancing the use of evolutionary strategies in computational problem-solving.

The paper by [4] describes the foundations of GAs, including implementation of selection, crossover operators, mutation and edge recombination crossover (ERX) techniques. This paper also explains different selection mechanism, like, roulette wheel selection, rank-based selection and tournament selection, which present a view on how to retain the best individuals while maintaining the population diversity. It introduced the building blocks hypothesis, which suggests that good solutions are formed by combining smaller, high-quality sub-solutions. The research introduced common crossover operators used for permutation problems, including partially matched crossover (PMX), order crossover (OX), and cycle crossover (CX). It describes mutation as a small, random change to an individual that ensures variation and prevents from premature convergence. Edge recombination crossover (ERX) is highlighted for preserving adjacency relationships between cities in TSP, which is important for maintaining good partial solutions across generations. The adaptive mutation strategy is discussed on how the rates of mutation can be varied during different stages of the problem-solving algorithm in order to increase the exploration and exploitation.

2) History of Simulated Annealing Algorithm: The paper [5] introduced the concept and implementation of Simulated Annealing as an optimization method, inspired by the proccess of annealing, where the metals are heated at a certain temperature and then cooled down to achieve a low-energy state. The algorithm works on ocassionally accepting worse-case solution to escape the local minima, with the chances of accepting such solutions decreasing as the algorithm progresses. The paper talks about working on several applications including TSP, integrated circuit design and graph partitioning. SA has emerged as a versatile method as it avoid being trapped in local minima making it suitable for solving problems with complex solution spaces.

The paper [6] independently introduced Monte Carlo simulation as a novel algorithm by drawing an analogy with statistical thermodynamics. The algorithm simulates the annealing process and in this context, the energy corresponds to the route length, and the temperature controls the chances of accepting route changes that might increase its length initially, thus making it easy for the algorithm to escape local minima. The methods works on the Boltzmann-Gibbs distribution principles and it begins with the random permutation of cities to generate new generation iteratively to improve the solution and then decides whether to accept them based on the Boltzmann-Gibbs distribution. The author was able to conclude that the algorithm can find solution, close to or equal to the optimal route. The paper acknowledges the gaps in theoretical convergence results and highlights the practical effectiveness of this thermodynamical approach.

III. Genetic Algorithm

A. Basic Genetic Algorithm

Genetic Algorithm is a metaheuristic optimization algorithm that performs on natural selection. It begins with generating a population of random routes, where each route is likely to be a possible solution. The algorithm calculates the total distance to find the effectiveness of the generated route, and the shortest distance is considered as better result. The routes are calculated based on their fitness level, so routes are ranked accordingly and the routes with short distances are considered a better pick for reproduction and then, Selection is performed using a combination of elitism and roulette wheel selection [4]. Elitism is used in preserving the top-performing routes and roulette wheel selection probabilistically selects routes based on their fitness level. From all of the selected routes, two of the routes are considered parents and are connected using a technique namely ordered crossover to produce new offspring. This method then takes a segment from one parent and keeps filling in the rest of the cities or points from other parents, keeping their contact intact. During the execution of algorithms, there are high chances of the algorithm getting stuck, so a mutation step is added. Mutation swaps the two cities in a route with a low probability, that helps to maintain genetic diversity. The process of selection, crossover, and mutation [2] repeats for a set number of generations and, in each generation, the process keeps preserving the best routes in each generation and continues until the algorithm finds a better solution.

B. Advanced Genetic Algorithm

The advanced Genetic Algorithm (GA) enhanced the basic GA version by integrating more complicated techniques for crossover and mutation. The main improvements are the implementation of the Edge Recombination Crossover (ERX) [4] method and the adaptive mutation rate. ERX assists in retaining adjacency information from both parent routes. To maintain the integrity of key route structures, this method results in new routes with greater uniformity. In another method called adaptive mutation rate, mutation decreases as generations go on. This approach starts with a high mutation rate in order to explore a wider range of solutions and slowly reduces it to focus more on improving the best solutions found. These improvements carry out the search for better solutions more efficient compared to basic genetic algorithm especially on larger or more complex instances of the TSP problem.

In this advanced GA, the creation of the initial population is similar to the basic GA. The enhancement is the use of ERX method to generate offspring by combining the edges from both parent routes, promising that significant structures are preserved. The adaptive mutation rate introduces variations by swapping cities with a probability that reduce over time, keeping a balance between exploration and exploitation. This method iterates for 500 generations, finally returning the best route found and can be easily visualized and analyzed.

IV. Simulated Annealing

A. Basic Simulated Annealing

Simulated Annealing (SA) is another metaheuristic optimization technique, which is inspired by the annealing process in metallurgy [5]. The metallurgy process helps to remove defects, making them more ductile and reducing the minimal energy from the material. The algorithm begins at an initial high temperature with the randomly generated routes . The total distance of the routes is calculated in advance to estimate the approximate cost. To generate neighboring solutions, the algorithm performs iteration by swapping two cities in the current route. SA works on the principle of always accepting routes with a shorter distance. If the newly generated route is longer, it may still be accepted based on the chances that it will decrease with the temperature. The acceptance probability function is implemented to determine this probability, which uses an exponential function of the difference in cost and current temperature. By accepting the worst results at an early stage, this mechanism allows the algorithm to escape the local minima, thereby leaving scope for wider solutions. As the algorithm progresses, the temperature automatically keeps decreasing, and it becomes less likely to accept the worse solution, leaving a wider space for the best solutions. The process repeats until the temperature reaches very low or maximum number of iterations. The final result is considered as the best solution.

B. Advanced Simulated Annealing

This is the enhanced version of basic SA that takes some improvements about result and convergence number. This is done by varying the perturbation strength and utilizing a local search technique (2-opt algorithm). Perturbation strength is used to determine the number of cities to swap when generating neighboring solutions. Instead of swapping just 2 cities at a time, this method allows swapping large numbers of cites that will shuffle routes significantly and may help in better exploration for potential good solutions. In addition, the 2-opt algorithm is applied for 100 iterations at a regular interval to the best route identified. The 2-opt algorithm works by eliminating two edges from the route iteratively and reconnect them in a different way to minimize the tour length, thus improving the optimization process. The temperature is then systematically reduced following a cooling schedule with a cooling rate of 0.99. This gradual reduction in the temperature will provide the algorithm with sufficient time to explore the solution space. Similar to basic SA, new solutions are accepted based on the probabilistic approach. The periodic application further refines the best route found. This hybrid method of combining global search with local search boosts the algorithm's ability to discover high-quality solutions significantly.

V. Performance

This section talks about the performance evaluation used for solving the TSP. Considering the Integer Programming, Genetic Algorithm, and Simulated Annealing, the solution covers the preciseness of the results in terms of distance and computational time required. The execution time keeps changing slightly. The seed value of last 4 digit of student ID is provided, so the process produce same result each time.

A. Integer Programming

IP algorithm was implemented to find the shortest distance for a roundtrip of 29, 40, and 80 cities of Bavaria and it consistently produced shortest possible route showcasing it capability to find the optimal solution. For a round trip of 29 cities, it successfully determined a route of 9074.15 km in just 3.52 seconds. As the problem complexity was increased and set for a roundtrip of 40 cities, the computational time increased to 4.79 seconds, resulting in a route of 9457.43 km. Further increasing the roundtrip of 80 cities, the algorithm witnessed the computational time rise to 10.06 seconds, with the route length extending to 13701.54 km. The results show the ability of IP to efficiently find optimal solutions for small-sized instances, matching well with the literature that recognizes IP for its exactness in solving combinatorial optimization problems like TSP. Figure 1 display the graph for the roundtrip of cities for IP.

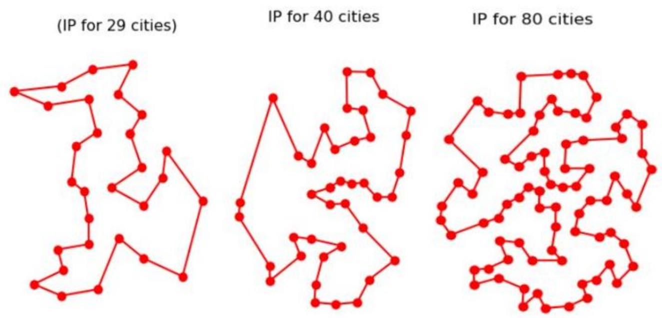

Fig. 1. IP roundtrip of 29,40 and 80 cities

B. Genetic Algorithm

The advanced Genetic Algorithm (GA) is built on the basic methodology set up by the basic GA by incorporating advanced techniques like Edge Recombination Crossover (ERX) and an adaptive mutation rate. ERX is particularly useful as it preserves adjacency information from both parents, thus creating offspring routes that are more coherent and logical. The adaptive mutation rate starts at a high rate to explore a wide solution space and decreases over time to fine-tune the best solutions. This advanced approach significantly improved the quality of solutions than the basic approach. For a roundtrip of 29 cities, the advanced GA found a route of 9200.52 km in 11.69 seconds. For 40 and 80 cities, the advanced GA found routes of 10775.64 km and 21884.39 km in 14.89 and 50.29 seconds, respectively. These results depicts the advanced GA's effectiveness in balancing exploration and exploitation, as supported by the literature. Figure 2 depicts the graph for the roundtrip of cities for GA.

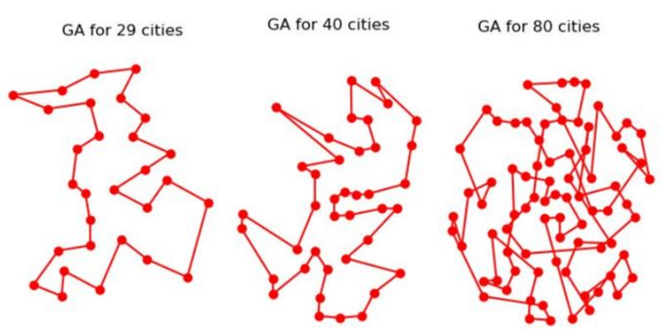

Fig. 2. GA roundtrip of 29, 40 and 80 cities

C. Simulated Annealing

Similar to the Genetic Algorithm, Simulated Annealing (SA) also enhances the basic SA by adding perturbation strength and the 2-opt local search technique. Perturbation strength assists in controlling the extent of changes in neighboring solutions, which allows for substantial changes and better exploration of the solution space. To further optimize the best route by reducing its distance, the 2-opt algorithm is applied at regular intervals. For the 29-city problem, the advanced SA found a route of 9905.16 km in 3.62 seconds. For a roundtrip of 40 and 80 cities, it found the routes of 10974.92 km and 15116.73 km in 11.38 and 98.99 seconds, respectively. This approach improved the performance of SA than the basic approach. These findings are consistent with the literature, which highlights SA's ability to escape local minima and effectively explore the solution space. Figure 3 depicts the graph for the roundtrip of cities for SA.

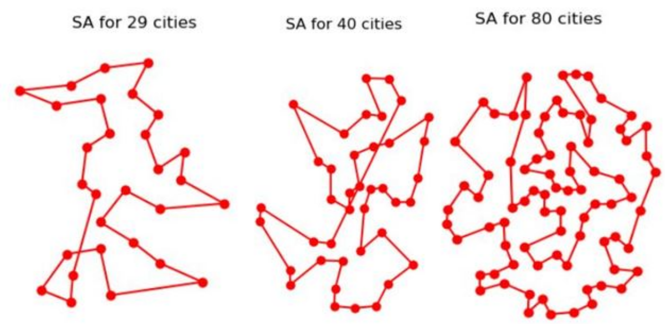

Fig. 3. SA roundtrip of 29,40 and 80 cities

VI. Conclusion

In the context of the project, we applied and compared the Integer Programme (IP), Genetic Algorithm (GA), and Simulated Annealing (SA) and their enhanced forms for solving TSP. The results indicate that different optimization approaches have their own advantages and trade-offs when dealing with the TSP. IP offered optimal solutions for relatively small TSP instances, however, it became computationally in-efficient working on larger dataset. Basic GA and SA provided faster solutions than IP, however, they sometimes reached suboptimal routes in the map. Advance features were incorporated in GA and SA to improve their performance. Advanced GA had features like, edge recombination and adaptive mutation, and the advanced SA, with adaptive cooling and 2-opt local search, that were found to improve the solution and scalability features. In general, the more sophisticated heuristic methodologies achieved better trade-offs between accuracy and time and might be better suited for larger actual multivariate optimization issues. Logistics industry are a great example of TSP, particularly in route optimization for delivery services and supply chain management.

References

[1] Alexander Schrijver. On the history of combinatorial optimization (till 1960). *Handbooks in operations research and management science*, 12:1–68, 2005.

[2] John Grefenstette, Rajeev Gopal, Brian Rosmaita, and Dirk Van Gucht. Genetic algorithms for the traveling salesman problem. In *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, pages 160–168, July 1985.

[3] George Dantzig, Ray Fulkerson, and Selmer Johnson. Solution of a large-scale traveling-salesman problem. *Journal of the operations research society of America*, 2(4):393–410, 1954.

[4] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA, 1989.

[5] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[6] Vladimír Černỳ. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications*, 45:41–51, 1985.