Analysis of Approximation Algorithms: Travelling Salesman Problem

Abstract

The Traveling Salesman (TSP) is the search for a minimum cost Hamiltonian circuit connecting a set of locations. There have been many ways devised to solve this problem which include branch and bound, greedy and dynamic programming first and later methods like Ant Colony Algorithm (ACA), Genetic Algorithm(GA), Particle Swarm Optimization (PSO) came into picture. The enumeration methods available now has been an effective method for solving smaller scale TSP problem, but to obtain the exact solution of larger scale TSP problem,it becomes very difficult. As a result, there have many papers put forward for solving TSP problem with approximate solution.
On 1976 N.Christofides put foreword a heuristic algorithm for solving travelling salesman problem of n nodes which improved the approximation ratio by 50 percent than the previous algorithm i.e the 2 approximation algorithm. Over the years there has been many improvisations to the ratio provided by Christofides. In this paper we will discuss a few of them and how much improvement that was shown than the previous ones.

Here we came upon the another general form of the same problem which is known as travelling salesman Path Problem (TSPP) which aims to find a path from two points after traversing all the vertices of the graph exactly once. The best approximation ratio of Christofides-type approximation algorithm found for this is 5/3 rather than 3/2 which was provided by Hoogeveen's algorithm.

There has been some notable progress in approximating the graphic TSP and TSPP also by the likes of Shayan Oveis Gharan,Mömke and Svensson and Kleinberg and Shmoys who improved this ratio to (1+sqroot(5))/2 $<$ 1.618034.

Mömke and Svensson gave a better approximation ratio later which is 1.461 for graphic TSP which is factor of 3 sqroot(2) $+ \varepsilon =$ 1.586$+\varepsilon$ for graphic TSPP, for any $\varepsilon > 0$. Mömke and Svensson used matching in a different way. In this approach unlike previously they didn't add edges of a matching to a spanning tree to make it Eulerian,instead the matching edges are first added and removed. And this was guided by pairing of edges that are removable that encodes the information on which the edges are removed from the graph one by one without disconnecting it. Ratio of approximation was found to be 5/4 in this case.

Introduction

The Traveling Salesman Problem (TSP) is analogous to a real life problem.Consider a salesman who visits many cities for his job. Naturally, he would want to take the shortest route through all the cities.The problem is to find this shortest route without taking lots of computation time. The goal of metric TSP is to find shortest Hamiltonian circuit given cost on a set of vertices.
Christofides' first aimed to find a minimum-cost spanning tree S in the complete graph with vertex set V and then added a parity correction to minimum cost spanning tree S.

Basic Terms

Method Of Enumeration

This is a combinatorial solution.It evaluates of all possible sequences in the total number of $(n - 1)!$. This has its own merit like global optimum is always found.But, it is not advisable when number of nodes are high.With every added node the solution space increases exponentially and even the most powerful computers find it tough to provide optimum solution within reasonable time.

Dynamic Programming

The main limitation of the enumeration method is the need for high time and space requirement when the number of city increases.Dynamic Programming which is based on enumeration often tries to reduce the amount of enumeration by reducing those decision sequences which are either not optimal or from where there us no possibility of reaching the feasible solution. This like divide and conquer also divides the problems into subproblems except in dynamic programming they were overlapping where as in DnC, they were independent.

Greedy Algorithm

This optimisation strategy tries to solve the problems using some sequence of decisions.These set of decisions are made in a stepwise manner based on local option available at every step. It only gives a suboptimal solution and works only for complete graphs. Like Kruskal's algorithm,it first sorts the edges in the increasing order of weights.and then starting with the least cost edge, look at the edges one by one and select an edge only if the edge, does not cause a vertex to have degree three or more and does not form a cycle, unless the number of selected edges equals the number of vertices in the graph.

Branch and Bound

The word, Branch and Bound refers to all the state space search methods in which we generate the children of all the expanded nodes, before making any live node as an expanded one. In this method, we find the most promising node and expand it. The term promising node means, choosing a node that can expand and give us an optimal solution. We start from the root and expand the tree until unless we approach an optimal solution.

Ant Colony Optimization

The ant system was First introduced by Marco Dorigo in 1992 and was Result of research on computational intelligence approaches to combinatorial optimization.

Ant Colony Optimization (ACO) studies artificial systems that take inspiration from the behavior of real ant colonies and which are used to solve optimization problems.As ants are Almost blind and Incapable of achieving complex tasks alone, they Rely on the phenomena of swarm intelligence for survival.By Using stigmergic communication via pheromone trails they are Capable of establishing shortest-route paths from their colony to feeding sources and back. The ants Follow existing pheromone trails with high probability. So,the more ants follow a trail, the more attractive that trail becomes for being followed.

Genetic Algorithm

Genetic algorithm (GA) as a computational intelligence method is a search technique used in computer science to find approximate solutions to combinatorial optimization problems. The genetic algorithms are more appropriately said to be an optimization technique based on natural evolution. They include the survival of the fittest idea algorithm. The idea is to first guess the solutions and then combining the fittest solution to create a new generation of solutions which should be better than the previous generation.

Motivation

Since TSP is a problem that has so much real world applications,the need to work on it was felt.Also Its very important problem in AI and Research domain. Sometimes information available in real life system is of vague, imprecise and uncertain nature.So,it becomes even more difficult to formulate a particular mathematical formulae to solve this.Chrostofides' algorithm which has been present for over 45 years still proviPaperes better performance guarentee for metric TSP.So, reading analysing and improving an approximation algorithm by Christofides which has been present for so long is itself motivating.

Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem

Report Date:1976-02-01
This paper provides us a new heuristic approach of solving the travelling salesman problem. Rw is the worst case ratio of the value of the answer obtained a heuristic approach to the value of the optimal solution Prior to this paper Garey and Johnson in their paper "Complexity of near optimal graph coloring"showed that to find a polynomial growth graph coloring algorithm with Rw < 2 is as hard as to find polynomial graph for hard coloring. Later Johnson came up with an algorithm of Rw < 11/9 was for the loading and packaging problem. This paper proposed a method with Rw < 3/2, which was a 50 percentage improvement over the last best known value of Rw < 2.

Algorithm and complexity :

The algorithm proposed in the paper can be divided into 2 major parts or steps:
1.Calculation of a shortest spanning tree.
2.Finding a minimum-cost perfect matching.
Complexity of the first step is O(n2) whereas the second step can be performed in the O(n3) complexity. Hence, the overall complexity of the new heuristic algorithm provided in this paper is O(n3)

Approximation using 2-RNN.

Before knowing 2-RNN its necessary to know about K-RNN.
k-RNN algorithm:
For every combination of the k vertices v1, v2,... vk , a partial tour T = (v1, v2,... vk) is created and the vertices v1,v2.....vk are marked as visited.
i is then set as k.
vi+1 is selected as the nearest unvisited neighbor of vi and vi+1is then appended to T.
There is a possibility that there can be multiple nearest neighbours. If that happens any of them can be selected and then vi+1 marked as visited and i is incremented by 1.
Based on the above we find n!/(n-k)!tours. Among them the shortest is selected as result.
2-RNN is k-RNN with k=2.
It can seem like nearest neighbor algorithm. But the catch here is that we start from edge unlike the former which starts from node. Its based on the fact that for every set of points in the plane, there exists a degree-5 MST. And we can easily convert a degree-5 MST to a spanning tree of degree 4 by finding all path which conect a subset of nodes and finding the shortet path and taking the union of them.
1. Root the MST at a leaf r.

2. For each vertex v ϵ V
Compute the shortest path P visiting v and all its children.
3. Return T4, the tree formed by the union of the paths P
By comparing the above algorithm and 2 RNN algorithm we conclude that this goes for local optimisation unlike 2RNN which goes for global optimisation.And global optimisation approach outperforms local for non decomposible problems like that of TSP.

The Salesman's Improved Paths: 3/2 + 1/34 Integrality Gap and Approximation Ratio

In this paper an improved analysis for the metric st path TSP is given by a polynomial time algorithm which performs better than the subtour elimination LP by 1.53 times. It considers the very idea that edge deletion leads to disconnected tress and uses parity correction to reconnect the tree partly after deleting some edges in Christofides' tree.

In this paper they made slight changes to the best-of-many Christofides' algorithm. First some edges are deleted from the spanning tree with the optimism that the resultant forest to be reconnected during parity correction. But in worst case scenario when it does not get reconnected they predict the cost of reconnection, and take the help of LP for having random choices for reconnecting edges. For reducing the complexity of LP they have used bipartite machings or netowork flows which results in decrease in time complexity to polynomial time unlike the subtour elimination LP.

They also took the help of the convex combination which was given by Gottschalk-Vygen's to have an upper bound the cost of reconnection. To have polynomial time complexity for the the convex combination they altered it using matroid partition.

Performance Analysis

Over the years there has been many attempts to better the approximation algo given by Christofides.
While Christofides' algorithm which is of type symmetric TSP has an approximation of 3/2 and time complexity of O(n^3).
2RNN algorithm which is again based on symmetric TSP has the ratio of 5/4 but the time complexity here is more than that of Christofides' i.e O(n^4).
Random sampling algorithm has 3/2- ε as ratio.
By using Novel use of matching, the ratio of 13/9 is found. Finding a cycle cover with relatively few cycles for cubic bipartite graph take polynomial time and has 9/7 as approximate ratio.
Another approximation algorithm with polynomial time complexity named Ear-decomposition optimization using forest representations of hyper graphs gives the ration as 7/5.

For metric TSP by consecutive path cover improvement 8/7 ratio is found. Here we came upon the another general form of the same problem which is known as travelling salesman Path Problem (TSPP) which aims to find a path from two points after traversing all the vertices of the graph exactly once. The best bound of this is 5/2 which was first provided by Hoogeveen's algorithm. In Graphic TSP, our aim is to find a minimum cost circuit visiting all the nodes at least once. So, we can apply the same logic as above to Graphic TSPP.

There has been some notable progress in approximating the graphic TSP and TSPP in recent times. Shayan Oveis Gharan obtained an approximation of 3/2- ε for Graphic TSP. In this case first an optimal solution of LP relaxation is computed. Then the solution is expressed as lambda uniform distribution of spanning trees. And finally, sampling of a Spanning Tree T from this distribution is done and after that a minimum cost matching on odd degree vertices of T is added.

Conclusion

In this report we read and try to provide overview of different approaches used for as approximation for travelling salesman problem. Though the results of these papers are promising, they are just the beginning. In future even better results will be found with an improved approximation ratio. Further research can be done to improve the time complexity of the algorithm from O (n^3).

References

1. N. Christofides. 1976. Worst-case analysis of a new heuristic for the travelling salesman problem. Tech. Rep. 388, Graduate School of Industrial Administration, Carnegie Mellon University

2. Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. 2011. A randomized rounding approach to the traveling salesman problem. In Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS'11). 550–559.

3. A. Sebo and A. Van Zuylen, "The Salesman's Improved Paths: A 3/2+1/34 Approximation," 2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS), New Brunswick, NJ, 2016, pp. 118-127, doi: 10.1109/FOCS.2016.21.

4. J. A. Hoogeveen. 1991. Analysis of Christofides' heuristic: some paths are more difficult than cycles. Opera- tions Res. Lett. 10, 5, 291–295. DOI:http://dx.doi.org/10.1016/0167-6377(91)90016-I

5. Hyung-Chan An, Robert Kleinberg, and David B. Shmoys. 2015. Improving Christofides' Algorithm for the s-t Path TSP. J. ACM 62, 5, Article 34 (November 2015), 28 pages. DOI:https://doi.org/10.1145/2818310