TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 15/20 pp459-465
Volume 12, Number 4, August 2007

Two-Level Genetic Algorithm for Clustered Traveling Salesman Problem with Application in Large-Scale TSPs

DING Chao (丁 超), CHENG Ye (成 晔)[**], HE Miao (何 苗)

Department of Industrial Engineering, Tsinghua University, Beijing 100084, China

Abstract: Let $G = (V, E)$ be a complete undirected graph with vertex set V, edge set E, and edge weights $l(e)$ satisfying the triangle inequality. The vertex set V is partitioned into clusters $V_1, V_2, \ldots, V_k$. The clustered traveling salesman problem (CTSP) seeks to compute the shortest Hamiltonian tour that visits all the vertices, in which the vertices of each cluster are visited consecutively. A two-level genetic algorithm (TLGA) was developed for the problem, which favors neither intra-cluster paths nor inter-cluster paths, thus realized integrated evolutionary optimization for both levels of the CTSP. Results show that the algorithm is more effective than known algorithms. A large-scale traveling salesman problem (TSP) can be converted into a CTSP by clustering so that it can then be solved by the algorithm. Test results demonstrate that the clustering TLGA for large TSPs is more effective and efficient than the classical genetic algorithm.

Key words: clustered traveling salesman problem (CTSP); traveling salesman problem (TSP); Hamiltonian cycle; genetic algorithm; integrated evolutionary optimization

Introduction

Let $G = (V, E)$ be a complete undirected graph with vertex set V, edge set E, and edge weights $l(e)$ satisfying the triangle inequality. The vertex set V is partitioned into clusters $V_1, V_2, \ldots, V_k$. The clustered traveling salesman problem (CTSP) seeks to compute the shortest Hamiltonian tour that visits all the vertices, in which the vertices of each cluster are visited consecutively. There are many applications of the CTSP[1]. The CTSP is NP-hard, and the traveling salesman problem (TSP) can be viewed as a special case of the CTSP in which there is only one cluster or each cluster has only one vertex. There are several variants of the problem depending on whether the start and end vertices of a cluster have been specified. This paper focuses on the variant with unspecified end vertices.

Anily et al.[2] developed a 5/3-approximation algorithm for the ordered CTSP with a prespecified visiting sequence for the clusters. Guttmann-Beck et al.[3] proposed approximation algorithms with performance bounds for some variants of the CTSP. For the CTSP with unspecified end vertices, their algorithm first uses a modified Christofides' algorithm to get the shortest free ends Hamiltonian paths in each cluster. After the first step, the two end vertices for each cluster and the intra-cluster paths are specified. Then a rural postman problem algorithm is used to connect all the intra-cluster paths to form a whole tour. This algorithm favors the intra-cluster Hamiltonian paths, which implies that the inter-cluster paths may be sacrificed when the end vertices in each cluster are already determined.

Sheng et al.[4] proposed a two-level-TSP hierarchical algorithm favoring inter-cluster paths. They first specify the shortest inter-cluster paths connecting every cluster. The start and end vertices are then specified for each cluster. Then a modified Christofides' algorithm is used to get the shortest Hamiltonian paths with two

Received: 2006-09-30; revised: 2006-11-03
** To whom correspondence should be addressed.
E-mail: chengye@mail.tsinghua.edu.cn; Tel: 86-10-62784548

specified end vertices in each cluster. At last a whole tour is formed by combining the paths generated in both levels. Sheng et al.[4] also showed that the penalties caused by favoring the intra-cluster Hamiltonian paths and the inter-cluster paths are comparable.

This paper describes a two-level genetic algorithm (TLGA) for the CTSP that favors neither of the two levels of paths, thus avoiding the disadvantages of both algorithms. The general idea is to find the shortest Hamiltonian cycle for each cluster, to delete a selected edge for each cycle to form an intra-cluster path, and then to connect all the intra-cluster paths in a certain sequence to form a whole tour. In the lower level, a genetic algorithm (GA) is used to find the shortest Hamiltonian cycle rather than the shortest Hamiltonian path for each cluster. In the higher level, a modified genetic algorithm is designed to determine which edge will be deleted from the shortest Hamiltonian cycle for each cluster, and the visiting sequence of all the clusters with the objective of shortest traveling tour for the whole problem. Since the shortest Hamiltonian cycle rather than the shortest Hamiltonian path is formed in the lower level, it leaves open the question of choosing the edge to be deleted. The higher level algorithm has the freedom to choose the edge to be deleted for each cluster while searching for the shortest whole tour.

The clustering method can be used to convert a large-scale TSP into a CTSP which can then be solved by the TLGA. Test results demonstrate that the clustering TLGA algorithm for large-scale TSPs is more effective and efficient than the classical genetic algorithm. The increase in computing time of clustering TLGA is almost proportioned to the problem size.

1 Two-Level Genetic Algorithm for the CTSP

The ordinal encoding scheme[5] was used in the TLGA. Under this scheme, any vertex is assigned a unique unsigned integer from 1 to the number of vertices. The ordinal encoding scheme uses genes that are integers. Chromosomes which represent the traveling paths are permutations of these integer genes. A gene segment is defined as a permutation of the serial numbers of vertices in a cluster. Note that a chromosome is also a permutation of all the gene segments with one gene segment per cluster.

In the lower level, the genetic algorithm finds the (near) shortest Hamiltonian cycles for each cluster. In the higher level, a modified genetic algorithm is designed to choose the edge to be deleted from the shortest Hamiltonian cycle for each cluster, and simultaneously determine the visiting sequence for all the clusters with the objective of the shortest entire traveling tour.

1.1 Lower level genetic algorithm

The objective of the lower level genetic algorithm is to find the shortest Hamiltonian cycle for the given vertices in each cluster. This is a classical application of the genetic algorithm to the traveling salesman problem[6]. In most cases the genetic algorithm cannot assure that it will find the shortest Hamiltonian cycle, but it will find an acceptable cycle.

The genetic algorithm for the TSP is applied cluster by cluster. The result of the lower level genetic algorithm is tours $T_1, T_2, \ldots, T_k$ for clusters $V_1, V_2, \ldots, V_k$. A simple example is shown in Fig. 1.

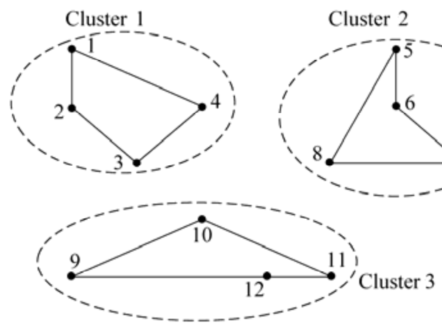

Fig. 1 Example result of the lower level GA

1.2 Higher level genetic algorithm

The objective of the higher level genetic algorithm is to form an entire tour that is as short as possible based on the tours generated in the lower level. In each generation, the designed genetic algorithm selects two adjacent vertices in tour T_i as end vertices for cluster V_i (that is to select the edge to be deleted from T_i), specifies which one will be the start vertex and which one will be the end vertex, and determines a sequence for visiting all the clusters.