A large-scale new variant of the capacitated clustering problem (VCCP)

CS357 Project Proposal, Autumn 2022, IIT Indore

By Amit Kumar Makkad, 200001003

Introduction

In this project, I will work to solve a problem encountered in a green energy investment project of the Champagne–Ardenne region of France. Here, the local government wants to build 4 plants in 6 potential sites to produce the ethanol from the wheat straw at the minimum total cost of opening plants and assigning suppliers to the plants. This problem (**VCCP**) is a new variant of the Capacitated Clustering Problem (**CCP**), which is a multi-objective problem with both equality and inequality constraints and it is one of the most widely studied location problems with various applications, for example in vehicle routing (Koskosidis & Powell, 1992) and political districting (Bozkaya, Erkut, & Laporte, 2003), etc. Here the main objective of our project will be to implement algorithms to solve the problem and to understand the idea behind implementation.

Problem statement

In the capacitated clustering problem (CCP), we have to partition a set of n customers with given demand into p clusters with the maximum capacity constraint on each cluster. Every customer must be assigned to exactly one cluster. For each given cluster, a vertex is defined as the cluster centre whose customers are all customers in the cluster. The objective is to find p optimal clusters and their corresponding cluster centres so that the sum of the distances from the cluster centres to their customers is minimized.

However, our problem (VCCP) is that we have to **minimize the cost of opening plants in some of potential sites and assigning suppliers to them**. Each plant is subject to a minimum capacity requirement, where each supplier is assigned to one plant at most and when a supplier is assigned to a facility, the former will supply its entire available volume to the latter. For an open plant, if its minimum capacity requirement has been attained, then assigning more suppliers to it will increase the cost. This case will never happen in the optimal solution of the problem. The total supply volume of all suppliers is greater than the total requirement of all plants.

Compared with the CCP, the VCCP has the following new features:

(1) Each open plant has a minimum capacity requirement that must be satisfied.

(2) Unlike the CCP where each customer has to be assigned to one cluster centre, in the VCCP, each supplier can be assigned to at most one plant.

These two features make our problem a variant of the CCP, denoted as VCCP in the following discussion. They cannot be transformed each other.

Mathematical Formulation

Let S be a set of suppliers with positive supply volume SV_i and P = {1,...,n} denote a set of potential sites for plants (facilities) with an associated minimum capacity requirement CR_j and a fixed opening cost FP_j and we have to open exactly p Plants. C_{ij} denotes the cost of assigning supplier i ∈ S to the plant located at site j ∈ P. We define a binary variable Y_j, which takes 1 if a plant is located at site j ∈ P and 0 otherwise. Binary variable X_{ij} equals 1 if supplier i ∈ S is assigned to plant located at site j ∈ P and 0 otherwise.

The VCCP can then be formulated as the following binary integer program:

Model P: $Z = \min \sum_{i \in S} \sum_{j \in P} C_{ij} . X_{ij} + \sum_{j \in P} FP_j . Y_j$ …(1)

St $\sum_{j \in P} X_{ij} \leq 1 \quad \forall i \in S$ …(2)

$\sum_{i \in S} SV_{ij} . X_{ij} \geq CR_j . Y_j \quad \forall j \in P$ …(3)

$\sum_{j \in P} Y_j = p \quad \forall j \in P$ …(4)

$X_{ij}, Y_j \in \{0,1\} \quad \forall i \in S, \forall j \in P$ …(5)

The objective function (1) is to minimize the total cost of assigning suppliers to plants and establishing such plants.
The assignment constraint (2) guarantee that each supplier is assigned to at most one plant.
The constraint (3) satisfies minimum supply requirement.
The constraint (4) ensures that p plants are exactly located.

Algorithm

The above problem is NP Hard as if one plant is located (p = 1) at one potential site (n = 1), it reduces to a knapsack problem. Therefore, Heuristics are the only practical techniques for handling large scale instances of this problem. There are many Heuristic and metaheuristic approaches have been proposed for the problem, for example, a column generation approach, by using hybrid scatter search with path relinking, a tabu search algorithm, a bionomic approach, genetic algorithm, etc. However, Lagrangian relaxation approaches also have been

successfully applied to CCP problems, such as the Capacitated Facility Location Problem (Cornuéjols, Sridharan, & Thizy, 1991), the Single Source Capacitated Facility Location Problem (Beasley, 1993; Pirkul, 1987), etc.

In this project, I will try to implement the **Lagrangian relaxation approach**, in which there are two phases of dual optimization, the subgradient deflection method in the first phase and the standard subgradient method in the second phase, to approximately solve the problem by relaxing the assignment constraints. In this approach, the best Lagrange multipliers and the best upper bound found in the first phase are used as inputs for the second phase. At each Lagrangian iteration, a feasible solution is constructed from the optimal solution of the Lagrangian relaxed problem by applying a greedy algorithm. Furthermore, at the end of the subgradient method, the best feasible solution found so far is improved by a simple tabu search algorithm.

Also, I will try to implement **Scatter Search**, which is a population-based algorithm that stores solutions in a so called reference set and constructs new solutions by combining existing ones. First, an initial solution is created and it is improved. Then taking this as seed solution, the set of diverse and new solution is created, which considered as a candidate solution for the reference set. Then the search is carried out in a loop until stopping criteria is met. In each iteration, it constructs set of the reference solutions combined by some combining methods like path relinking, etc.

References

A Lagrangian relaxation approach for a large scale new variant of capacitated clustering problem Zhen Yang , Haoxun Chen , Feng Chu , 2010 ScienceDirect