Algorithmic Methods for Mathematical Models
Course Project

A large company produces huge amounts of data and wants to store them as monthly backups. The company has O different regional offices and we know, for each office o, the amount of data d_o in Peta Bytes (1 PB = 10^6 GB) that it needs to store.

The company has not the necessary resources to store such amount of data, so it decides to hire external backup providers.

Due to latency constraints, every office can use only some of the backup centers. Given an office o and a center c, this information it is specified by parameter u_{co}, equal to 1 if such a connection is allowed.

For security reasons, every backed data needs to be stored in two different backup centers.

After studying several possibilities, several backup providers are chosen, resulting in C different backup centers, where every backup center c can store at most k_c PBs of data.

Using a backup center c has a fixed cost f_c and an additional cost for every PB it stores. The cost of storing data will be computed in P segments of decreasing cost that are the same for all backup centers. For every cost segment, s_p specifies the cost per PB stored and m_p the minimum stored data to apply cost segment p.

> *Example*: Let us assume three cost segments, where storing less than 7 PBs costs 5,000 EUR per PB, less than 12 PBs costs 3,000 EUR each, and storing 12 PBs or more costs 2,000 EUR each. Then, storing 14 PBs will cost 14*2,000 EUR.

We assume that the above parameters for cost and data are integers and that only integer amounts of PBs from every office can be stored.

In this project, we have been asked to elaborate a plan for the company. We have to find out which backup centers will be used and how many PBs from each office every backup center will store so that the total cost is minimized.

1 Work to be done

- Formally state the problem.
- Devise an integer linear programming model for the optimization problem and implement it in OPL.
- Because of the complexity of the optimization problem, heuristic algorithms are needed. We are considering both GRASP and BRKGA meta-heuristics. Implement your meta-heuristics in the programming language you prefer.
- Compare the performance of solving the model and the heuristics in terms of computation time and quality of the solutions. To that end, generate increasingly large problem instances until while solving the model takes no longer than 2h.
- Compare the performance of the two meta-heuristics in terms of solving time and quality of the solutions for even larger problem instances.

2 Report

Prepare a report (8-10 pages) including:

- Problem statement.
- Integer lineal model, including the definition of the sets and parameters, the model itself and a short description of the objective function and every constraint.
- For the meta-heuristics, the pseudo-code of the GRASP *constructive and local search phases* algorithms, the greedy function and the equation describing the RCL, and for the BRKGA heuristic, the chromosome structure and the pseudo-code of the *decoder* algorithm.
- Comparative results.
- Together with the report, you should also provide all sources and instructions on how to use them, so that results can be easily reproduced. If you implemented an instance generation, please provide it as well.