

This document is an attachment of the paper: *Claudia Archetti, Francesco Carrabs, Raffaele Cerulli, **The Set Orienteering Problem***, published on European Journal of Operational Research 2018.

We document the file format of the instances. Files can be downloaded at <http://www.dipmat2.unisa.it/people/carrabs/www/DataSet/InstancesSOP.zip>

### ***Instances***

Instances are obtained by adapting instances for the Generalized Traveling Salesman Problem (GTSP) proposed in:

*“M. Fischetti, J. J. Salazar Gonzalez, P. Toth, **A branch-and-cut algorithm for the symmetric generalized traveling salesman problem**, Operations Research 45 (3) (1997) 378–394”.*

In particular, among all these instances, we select the instances for which the distance is defined as the Euclidean distance between customers’ coordinates. These instances have a number of vertices ranging from 52 to 1084 and a number of clusters equal to ~ 20% of the number of vertices. They are 51 in total (see the paper for more details about the instances generation).

Files are organized as follows.

- The depot is the node 1. This node is the only one inside the cluster zero.
- File name is structured as follows: *<TSP-instance-name>T<percentage of best/optimal GTSP solution>p<#kind of generation of profits>.sop*

A second set of instances has the abbreviation “**RND**” inside the name. In this set the vertices are randomly assigned to the clusters.

- File format follows TSPLIB format with an additional GTSP\_SET\_SECTION: One line for each cluster of customers:  
*<cluster\_id> <cluster\_profit> list:<customer\_id>*