

OR-Library

J E Beasley

OR-Library is a collection of test data sets for a variety of OR problems.

A full list of the test data sets available in OR-Library can be found [here](#).

Set covering

There are currently 87 data files.

50 of these data files are the test problem sets 4 to 6 and A to E from J.E.Beasley "An algorithm for set covering problems" European Journal of Operational Research 31 (1987) 85-93.

Problem sets 4, 5 and 6 are originally from the Balas and Ho set covering paper.

The following table gives the relationship between these test problem sets and the appropriate files:

Problem set	Files
4	scp41, ..., scp410
5	scp51, ..., scp510
6	scp61, ..., scp65
A	scpa1, ..., scpa5
B	scpb1, ..., scpb5
C	scpc1, ..., scpc5
D	scpd1, ..., scpd5
E	scpe1, ..., scpe5

20 of these data files are the test problem sets E to H from J.E.Beasley "A lagrangian heuristic for set-covering problems" Naval Research Logistics 37 (1990) 151-164.

The following table gives the relationship between these test problem sets and the appropriate files:

Problem set	Files
E	scpnre1, ..., scpnre5
F	scpnrf1, ..., scpnrf5
G	scpnrg1, ..., scpnrg5
H	scpnrh1, ..., scpnrh5

10 of these data files are unicast problems from the paper "Computational Experience with Approximation Algorithms for the Set Covering Problem", by T. Grossman and A. Wool, European Journal of Operational Research 101(1) pages 81-92 (1997).

The files scpcyc06, ..., scpcyc11 correspond to the CYC set of problems (number of edges required to hit every 4-cycle in a hypercube). The files scpcclr10, ..., scpcclr13 correspond to the CLR set of problems (number of 4-tuples forming the smallest non-bi-chromatic hypergraph). These files have been contributed by A. Wool

The format of all of these 80 data files is:
 number of rows (m), number of columns (n)
 the cost of each column $c(j), j=1, \dots, n$
 for each row i ($i=1, \dots, m$): the number of columns which cover
 row i followed by a list of the columns which cover row i

For the files associated with the European Journal of
 Operational Research paper by Beasley:

- (a) the value of the optimal solution for each of these
 data files is given in the paper
- (b) the largest file is scpd5 of size 420Kb (approximately)
- (c) the entire set of files is of size 5800Kb (approximately)

For the files associated with the Naval Research Logistics
 paper:

- (a) heuristic solution values for each of these data
 files are given in the paper
- (b) the largest file is scpnrh5 of size 2600Kb (approximately)

There are also 7 data files associated with real-world set
 covering problems. These data files are:

rail507 with 507 rows and 63,009 columns
 rail516 with 516 rows and 47,311 columns
 rail582 with 582 rows and 55,515 columns
 rail2536 with 2536 rows and 1,081,841 columns
 rail2586 with 2586 rows and 920,683 columns
 rail4284 with 4284 rows and 1,092,610 columns
 rail4872 with 4872 rows and 968,672 columns

These data files arise from an application in Italian railways
 and have been contributed by Paolo Nobile.

As might be expected these problems have a number of special
 characteristics, specifically:

- (a) all column costs are either one or two
- (b) a column covers at most 12 rows
- (c) substantial reductions can be made by applying known
 row/column reduction tests

The format of these test problems is:
 number of rows (m), number of columns (n)
 for each column j ($j=1, \dots, n$): the cost of the column, the
 number of rows that it covers followed by a list of the rows
 that it covers

The largest file is rail4284 of size 61Mb (approximately).
 The entire set of files is of size 260Mb (approximately).

Click [here](#) to access these files

OTHER SOURCES

A number of set covering problems derived from Steiner triple systems can be found [here](#)