



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Large-scale test data set for location problems

Matej Cebecauer^{a,*}, Ľuboš Buzna^{b,c}^a Department of Transport Science, KTH Royal Institute of Technology, Teknikringen 10, SE-100 44 Stockholm, Sweden^b Department of Mathematical Methods and Operations Research, University of Žilina, Univerzitná 8215/1, SK-010 26 Žilina, Slovakia^c ERA chair for Intelligent Transport Systems, University of Žilina, Univerzitná 8215/1, SK-010 26 Žilina, Slovakia

ARTICLE INFO

Article history:

Received 26 July 2017

Received in revised form

13 December 2017

Accepted 4 January 2018

Available online 10 January 2018

ABSTRACT

Designers of location algorithms share test data sets (benchmarks) to be able to compare performance of newly developed algorithms. In previous decades, the availability of locational data was limited. Big data has revolutionised the amount and detail of information available about human activities and the environment. It is expected that integration of big data into location analysis will increase the resolution and precision of input data. Consequently, the size of solved problems will significantly increase the demand on the development of algorithms that will be able to solve such problems. Accessibility of realistic large scale test data sets, with the number of demands points above 100,000, is very limited. The presented data set covers entire area of Slovakia and consists of the graph of the road network and almost 700,000 connected demand points. The population of 5.5 million inhabitants is allocated to the locations of demand points considering the residential population grid to estimate the size of the demand. The resolution of demand point locations is 100 m. With this article the test data is made publicly available to enable other researches to investigate their algorithms. The second area of its utilisation is the design of methods to eliminate aggregation errors that are usually present when considering location problems of such size. The data set is related to two research articles: “A Versatile Adaptive Aggregation Framework for Spatially Large Discrete Location-Allocation Problem” (Cebecauer and Buzna, 2017) [1] and

DOI of original article: <https://doi.org/10.1016/j.cie.2017.07.022>

* Corresponding author.

E-mail address: matejc@kth.se (M. Cebecauer).<https://doi.org/10.1016/j.dib.2018.01.008>2352-3409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

“Effects of demand estimates on the evaluation and optimality of service centre locations” (Cebecauer et al., 2016) [2].

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	Applied mathematics, Operations research, Discrete optimization
More specific subject area	Location analysis, Geographic information systems
Type of data	graph of the road network, weighted demand points derived from GIS data and residential population grid
How data was acquired	Data set was created by combing publicly available data sets such as OpenStreetMap and residential population grid.
Data format	csv text files, shapefiles
Data source location	Slovakia (Longitude 17.001–22.110, Latitude 47.732–49.586)
Data accessibility	The data are available with this article. Moreover, data is published on the professional web page of one of the co-authors: http://frdsa.uniza.sk/~buzna/page5/page5.html

Value of the Data

- Data set can be used as a benchmark to design and experiment with new location algorithms intended to solve large-scale locational and spatial problems.
- Data set is applicable in the design and studies of new aggregation methods to minimise the impact of aggregation errors on the outcome of optimisation.
- Data set can be used to derive large number of medium and small size benchmarks by selecting specific geographic areas.
- Data set enables visualisation of results of optimisation algorithms in GIS.

1. Data

Central component of the benchmark Slovakia is the graph consisting of 1,956,067 georeferenced nodes, further defining 2,080,694 edges, representing the road sections covering the entire area of Slovakia. Some of these nodes (663,203) identify the potential population demand distribution derived from the residential population density. In the literature it is common to refer to these points as to demand points (DPs). A potential demand is located in the populated area approximately each 100 m and connected to the road network (see Fig. 1 for illustration).

With this article we deliver four different benchmarks (see Table 1), the main benchmark Slovakia and three benchmarks, Žilina, Košice and Partizánske. Benchmarks Žilina, Košice and Partizánske has been derived from the benchmark Slovakia by restricting the scope to selected geographical areas. All benchmarks have weights derived from the residential population grid [4].

Download English Version:

<https://daneshyari.com/en/article/6597027>

Download Persian Version:

<https://daneshyari.com/article/6597027>

[Daneshyari.com](https://daneshyari.com)