



## Possibilities of Solving Big GIS Projects on Desktop Computers

Dalibor BARTONĚK

Brno University of Technology, Faculty of Civil Engineering; Veveří 331/95, 602 00 Brno, Czech Republic  
European Polytechnic Institute, Osvobození 699, 686 04 Kunovice, Czech Republic  
bartonek.d@fce.vutbr.cz, bartonek.d@edukomplex.cz

### Abstract

We are witnessing great developments in digital information technologies. The situation encroaches on spatial data, which contain both attributive and localization features, and this determines their position unequally within an obligatory coordinate system. These changes have resulted in the rapid growth of digital data, significantly supported by technical advances regarding the devices which produce them. As technology for making spatial data advances, methods and software for big data processing are falling behind. Paradoxically, only about 2% of the total volume of data is actually used. Big data processing often requires high computation performance hardware and software. Only a few users possess the appropriate information infrastructure. The proportion of processed data would improve if big data could be processed by ordinary users. In geographical information systems (GIS), these problems arise when solving projects related to extensive territory or considerable secondary complexity, which require big data processing. The following methods are proposed to solve extensive projects in GIS:

1) For a large territory, it is recommended to split input data into segments relatively independent on each other and to make use of parallelism in further processing (this method was adopted in the project "Classification of the gas-pipeline surface area in the Czech Republic".

If this division is not feasible, the whole set of input data (e.g. in the case of network analysis) can be used.

2) Purpose-built reduction or optimization of input data by

a) data filtration (this method has been used in the case of laser scanning)

b) reducing the dimension of the problem, i.e. solving the task in a lower dimension and then reverting to a higher dimension. This method has been used when searching for an optimal route in a field with the RTK measurement method.

In order to gain a qualified estimation of relevant parameters (time solution, capacity of sources and structure of the research team), both methods require testing (pilot project).

Keywords: geographic information system (GIS), big data, parallel computing, classification, optimization, filtration, point cloud