Máté Cserép

EIT Digital Budapest DTC

PhD topic: Examination of methodological problems of vector and topological data models

PhD supervisor: Dr. István Elek, ELTE

Contact: mcserep@outlook.com

'The EIT Digital Doctoral School offers unique opportunities of geographical and industrial mobility to me, both as an integral part of its doctoral programme and in the forms of summer schools, workshops and symposiums organized by EIT Digital.'

Achievements & further plans

I am in the final year of my PhD studies at ELTE, with my doctoral topic being focused on analysing the methodological problems of vector and topological data models. Tightly combined with my doctoral research, I participate in an FP7 project named IQmulus whose main objective is to enable the use of large, heterogeneous **geo-spatial data sets** for better decision making through a **high-volume fusion and analysis information management platform**. Currently, I am planning to complete my geographical mobility at TU Delft in the spring semester of 2016, under the supervision of Roderik Lindenbergh, where I intend to research data analysis and change detection based on airborne laser scanning.

As almost every industry or administrative organisation has to deal with spatial data in some form, Geographical Information Systems (GIS) are moving from the scientific and public service area towards the private enterprise sector, resulting in an exponential growth in the amount of spatial data maintained and services developed.

Educational status at Spring semester of 2016:

RA	OR	BMD	GH	Mobility	BDExp.
✔	✔	✔	⚙	⌛	⚙

Reserach topic

The most extensively used spatial data formats do not contain any topological information, neither do they support any change detection or version management natively. Without thorough topological analysis, data imported from such formats is abound with joint errors and superfluous information, while storage space efficiency for multiple revisions also poses an unresolved issue.

As part of my doctoral research topic I aim to provide vector data structures that are capable of representing topological information without redundancy, while supporting the known spatial relations. The topological analysis of large data sets and the creation of such representations require high computational capacity, hence distributed computing and parallelization also have to be utilized to accelerate the progress. While version management is a common concept and tool in software development, it is rarely used in the field of Geographic Information Systems (GIS), though it could reduce the number of irreversible errors, significantly ease the merge of concurrent modifications and provide semantic information regarding the changesets. Current revision control software focus primarily on textual data as no method is available for the effcient storage of modifcations on any binary content.