

Oberseminar Theoretische Informatik

Wintersemester 2009/2010

Johannes Uhlmann

Exact Algorithms and Experiments for Hierarchical Tree Clustering

Montag, 23.11.2009 14:00 (c.t.) Seminarraum 3319 (Ernst-Abbe-Platz 2, 3.
Stock).

We perform new theoretical as well as first-time experimental studies for the NP-hard problem to find a closest ultrametric for given dissimilarity data on pairs. This is a central problem in the context of hierarchical clustering, where so far only polynomial-time approximation algorithms were known. By way of contrast, we develop efficient preprocessing techniques (known as kernelization in parameterized algorithmics) with provable performance guarantees and a simple search tree algorithm. These are used to find optimal solutions. Our experiments with synthetic and real-world biological data show the effectiveness of our algorithms and also demonstrate that an approximation algorithm due to Ailon and Charikar [FOCS 2005] often gives (almost) optimal solutions.

Homepage:

<http://theinfl.informatik.uni-jena.de/teaching/ws0910/oberseminar-ws0910>