



FP7-PEOPLE-2009-IRSES:  
Project ID 246647

**Optimization and its Applications  
in Learning and Industry  
(OptALI)**

IRSES

**Ongoing Deliverable D1.2**

**Description of Research Seminar:  
Locating a median strip in the plane**

Start date of the Workpackage: December 2010

Duration: 48 months

Due date of deliverable: November 2014

Actual submission date: April 2013

Participants: UGOE  
UNIKL  
DTU  
UOA  
UC

Author of deliverable: R. Schieweck ([r.schieweck@math.uni-goettingen.de](mailto:r.schieweck@math.uni-goettingen.de))

# Research Seminar

offered by Robert Schieweck (ES-UGOE-6)

in March 2013

in Auckland, New Zealand

**Subject:** Locating a median strip in the plane

**Problem:** In statistical linear regression the observed data points are sometimes subject to latent factors. If there are finitely many values for the latent factors this results in the fact that data points fall into classes which are unknown a priori and must be estimated simultaneously along with the usual regression parameters. From a geometrical point of view this amounts to locating a number of lines to approximate a given point set. This is an NP hard problem under various meaningful distance measures to determine the deviations from each data point to the lines and also for various ways of aggregating point-line deviations into a single objective function. We are considering the median objective which aims at minimizing the sum of norm-induced distances from each point to its respective closest line. Even though this problem can be solved by mixed-integer programming for certain distance measures the running times are discouraging. Therefore new methods are needed to solve this problem and the problem of locating a number of strips to approximate a given point set arises as a subproblem during the computation of lower bounds on the objective value which can then be used to discard large sets of candidate solutions for the original problem. We investigate the problem of locating a single strip in order to minimize the sum of distances to some given points.

**Main Results:** Since a strip is – informally speaking – a line with a certain thickness a natural question is whether classical properties for line location problems can be generalized to strip location problems. We show that

indeed for the problem of locating a single strip the *pseudo-halving property* for median lines also holds for strips and that the *weak incidence property* can be slightly altered to be applicable in the new setting. By using these properties we identify a finite candidate set for the median strip location problem which gives rise to a brute-force enumeration scheme for its solution in  $O(n^3)$  to  $O(n^4)$  worst-case running time. This approach can then be sped up by using a sweep line technique resulting in an  $O(n^2 \log n)$  time algorithm. The finite candidate set can also be used to find a set of multiple strips to approximate the given points. Computational results to evaluate the impact of the lower bounds generated by locating strips on solutions times for the original problem, i. e. locating lines, are pending.

Participants: students and researchers from UOA

Publication: -