14th Conference on

DATA ANALYSIS METHODS
for Software Systems

November 30 – December 2, 2023

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

Co-Chairmen:

Prof. Gintautas Dzemyda (Vilnius University, Lithuanian Academy of Sciences)
Dr. Saulius Maskeliūnas (Lithuanian Computer Society)

Programme Committee:

Dr. Jolita Bernatavičienė (Lithuania)
Prof. Juris Borzovs (Latvia)
Prof. Robertas Damaševičius (Lithuania)
Prof. Janis Grundspenkis (Latvia)
Prof. Janusz Kacprzyk (Poland)
Prof. Ignacy Kaliszewski (Poland)
Prof. Bożena Kostek (Poland)
Prof. Tomas Krilavičius (Lithuania)
Prof. Olga Kurasova (Lithuania)
Assoc. Prof. Tatiana Tchemisova (Portugal)
Prof. Julius Žilinskas (Lithuania)

Organizing Committee:

Dr. Jolita Bernatavičienė
Prof. Olga Kurasova
Assoc. Prof. Viktor Medvedev
Laima Paliulionienė
Assoc. Prof. Martynas Sabaliauskas
Prof. Povilas Treigys

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt
Tel. +370 5 2109315

Minimum-Sum-Of-Squares Clustering With (Net) Constraints for Cluster-Centres

Mindaugas Kepalas, Julius Žilinskas

Institute of Data Science and Digital Technologies
Vilnius University

mindaugas.kepalas@mif.stud.vu.lt

We present our research results on the problem indicated in the title. Minimum-sum-of-squares clustering is a famous data-science problem for which locally optimal solutions can be found by running the famous k-means algorithm. However, if one seeks to find the global (the best possible) solution of the problem, things become really challenging: this task in scientific literature is proven to be NP-complete. In our research, we modify the famous min-sum-clustering problem by introducing constraints for the placement of cluster centres: these must be placed in a subset of the space (as opposed to the original problem, where the centres can be placed anywhere in the space). In 2-dimensions, possible restrictions might be, for example, that the centres must be placed on a road (e.g., on a net), or cannot be placed in certain regions (e.g., lakes or forests). We report the results of our attempts to find the global solution to the presented problem.