TRUDY INSTITUTA MATEMATIKI I MEKHANIKI UrO RAN

Vol. 22 No. 3 **2016**

DOI: 10.21538/0134-4889-2016-22-3-144-152

MSC: 68W25, 68Q25

AN APPROXIMATION ALGORITHM FOR THE PROBLEM OF PARTITIONING A SEQUENCE INTO CLUSTERS WITH CONSTRAINTS ON THEIR CARDINALITIES

Received May 30, 2016

A. V. Kel'manov, L. V. Mikhailova, S. A. Khamidullin, V. I. Khandeev

We consider the problem of partitioning a finite sequence of points in Euclidean space into a given number of clusters (subsequences) minimizing the sum over all clusters of intracluster sums of squared distances from elements of the clusters to their centers. It is assumed that the center of one of the desired clusters is specified at the origin, while the centers of the other clusters are unknown. Very unknown cluster center is defined as the mean value of cluster elements. Additionally, there are a few structural constraints on the elements of the sequence that enter the clusters with unknown centers: (1) the concatenation of indices of elements of these clusters is an increasing sequence, (2) the difference between two consequent indices is bounded from below and above by prescribed constants, and (3) the total number of elements in these clusters is given as an input. It is shown that the problem is strongly NP-hard. A 2-approximation algorithm that is polynomial for a fixed number of clusters is proposed for this problem.

Keywords: partitioning, sequence, Euclidean space, minimum sum of squared distances, NP-hardness, approximation algorithm.

A. V. Kel'manov, Dr. Phys.-Math. Sci., Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia; Novosibirsk State University, Novosibirsk, 630090 Russia, e-mail: kelm@math.nsc.ru .

L. V. Mikhailova, Cand. Sci. (Phys.-Math.), Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia, e-mail: mikh@math.nsc.ru .

S. A. Khamidullin, Cand. Sci. (Phys.-Math.), Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia, e-mail: kham@math.nsc.ru .

V. I. Khandeev, doctoral student, Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia, e-mail: khandeev@math.nsc.ru .

Cite this article as:
A. V. Kel'manov, L. V. Mikhailova, S. A. Khamidullin, V. I. Khandeev, An approximation algorithm for the problem of partitioning a sequence into clusters with constraints on their cardinalities, *Trudy Inst. Mat. Mekh. UrO RAN*, 2016, vol. 22, no. 3, pp. 144–152 .