

**3RD INTERNATIONAL EURASIAN CONFERENCE
ON
MATHEMATICAL SCIENCES AND
APPLICATIONS**

BOOKS OF ABSTRACTS

**25-28 AUGUST 2014
VIENNA-AUSTRIA**

Parallel Clustering Algorithms Based on Mapreduce Hadoop

Madina Mansurova¹, Darkhan Akhmed-Zaki², Bolatzhan Kumalakov³, Bazargul Matkerim⁴

Abstract. This study represents an approach of application of parallel algorithms for data processing of Remote sensing. Hyperspectral information of Remote sensing finds application in many branches. Therefore problems of processing of hyperspectral information and decoding of space images are actual. The goal of this study is the implementation of parallel algorithms for clustering of hyperspectral images. In this study parallel clustering algorithms are based on MapReduce programming model and are implemented on the Hadoop platform. Research results are compared with the results with different cluster settings and MPI model.

We have designed an experiment to gain data on working implementations of three Hadoop solutions for clustering problem: ISODATA, ISODATA with combiners and clustering using Particle swarm optimization. The increasing of productivity is achieved by the process of data control. In this work the processed data are divided into two types: static, which are stored in the local file system, and state's data that participate in the exchange between the Map and Reduce functions. Iterative algorithms use specific scheme that reduces the amount of data participating in the exchange of MapReduce processes. The results presented in this study prove the effectiveness of MapReduce Hadoop technology for Remote Sensing data processing.

Keywords. MapReduce, Hadoop platform, MapReduce-based parallel clustering algorithms, Remote Sensing, ISODATA algorithm, Particle swarm algorithm.

References

- [1] Pughineanu C., Balan I. Parallel Algorithm Evaluation in the Image and Clustering Processing, Electronics and electrical engineering. System engineering, computer technology. – 2011. – No.4 (110), pp. 89-92.
- [2] Byung-I K. and Alan G. Parallel asynchronous particle swarm optimization, International Journal For Numerical Methods In Engineering. Vol. 67, pp. 578-595, 2006.

¹ Al-Farabi Kazakh National University, Almaty, Kazakhstan, mansurova01@mail.ru

² Al-Farabi Kazakh National University, Almaty, Kazakhstan, Darhan.Ahmed-Zaki@kaznu.kz

³ Al-Farabi Kazakh National University, Almaty, Kazakhstan, b.kumalakov@gmail.com

⁴ Al-Farabi Kazakh National University, Almaty, Kazakhstan, bazargulmm@gmail.com