

## **HIGH-PERFORMANCE DISTRIBUTED MODULE OF SPEAKER IDENTIFICATION SYSTEM LEARNING**

**A.I. Korolev, O.Y. Kudashev, D.M. Timoshenko**

*Speech Technology Center, Saint-Petersburg, Russia*

*a.nezachem@gmail.com*

An automatic speaker identification problem is quite a complex and time-consuming process. It is closely related to the methods of machine learning and mathematical statistics, as well as optimization theory and numerical computing. It's a fact of common knowledge that to achieve good recognition rate it's necessary to tune the recognition system to a particular phonogram base. The solution of this task implies processing of huge amounts of audio data and turns out to suit for being parallelized by high-performance distributed system very well.

In our report we would like to present some experience about constructing and using such a system as well as learning algorithms involved. The system itself is a cluster based on linux-driven independent computing nodes that perform parallel computations using both GPU and CPU and use MPI as inter-process communication means. As we expected, such approach turned out to be effective enough, and great performance gain was obtained.