

STSM Scientific Report

Reference: COST-STSM-IC1302-33931
STSM Applicant: Dr. Atanas Hristov, University of Information Science and Technology "St. Paul the Apostle". Ohrid, Macedonia
Host: Ass. Professor Georgios Petrou, Neapolis University, Cyprus
Period: 15.07.2016 - 22.07.2016
STSM Title: Multiobjective optimization for feature selection in text classification problems

1. Purpose of the Short Term Scientific Mission (STSM)

The purpose of the STSM was to enable Dr. Atanas Hristov to visit Ass. Professor Georgios Petrou at the Neapolis University in order to discuss his experience and to extend his knowledge in the area of high performance computing, more specifically in the area of Multiobjective optimization with main focus on the parallel multiobjective optimization algorithms.

In particular, my visit has: developed my skills as early-stage researcher through receiving immediate, face-to-face feedback on proposed work plan; expanded my network by introducing to experts at the School of Informatics at Neapolis University in Cyprus, as well as experts in the field of high performance computing, multi objective optimization, Social Networks, and Bid Data.

2. Description of the work carried out during the STSM

During the STSM, I spent one week at the School of Informatics at Neapolis University in Cyprus, under supervision of Ass. Professor Georgios Petrou. The work carried out during my visit has been mainly focused on identify and to analyze the general problems and modern trends in the area of high performance computing, multi objective optimization, and text classification problems, as well as, the tools and algorithms there are used by this research group in order to exploit the modern systems and platforms in efficient way. On the other hand we have very effective collaboration in the area of parallel data processing, more specifically in parallel implementation of multiobjective optimization algorithms, where we developed new parallel implementations of algorithm for multiobjective optimization for feature selection in text classification problems by using MPI and OpenMP parallel programming models. The proposed algorithm was evaluated and the parallel performance was tested on the computer cluster located at the High-Performance Computing Laboratory, School of Informatics, Neapolis University.

3. Description of the main results obtained

The main results have been conducted mainly together with my host Ass. Professor Georgios Petrou, and by support of his team. We managed to achieve all activities attached to the proposed work plan including:

1. Review, identify and analyze the general problems and modern trends in the area of multiobjective optimization and text classification problems with special focus on high performance algorithms, tools and application for this purpose.
2. Define the unsupervised feature selection problem as a multiobjective optimization problem and implement the base sequential procedure along with the set of various data sets used for the experiments.
3. Design a parallel evolutionary procedure for multiobjective optimization based on the island model, with different alternatives for exploring separate search spaces across different subpopulations (islands).
4. Implement the parallel procedure and evaluate the implemented search space strategies with the alternative objective formulations used to determine the quality of the text classification obtained with the selected features.

A novel multiobjective optimization algorithm for feature selection in text classification problems by using MPI and OpenMP parallel programming models has been developed, and his efficiently was evaluated on the computer cluster located at the High-Performance Computing Laboratory at School of Informatics. The experimental results have shown good parallel performance and efficiency of the proposed algorithms. The STSM also provided us with the opportunity to fine tune our research aims and objectives, commence a literature review of multiobjective optimization and text classification problems, and finalize our research timetable and intended output plans.

4. Future collaboration with host institution

We consider this STSM a very successful, and we estimate that the project will take several months to be completed. Since I have opened door for future collaborations with the group of Ass. Professor Georgios Petrou, we plan to collaborate with the host institution in the near future, in order to achieve two objectives: implementation and evaluation of the proposed algorithms on the cluster located at my home institution, and improvement of the disadvantages that appear during the evaluation of the algorithm.

5. Foreseen publications/articles resulting or to result from the STSM

Since our project is still in very early phase, we intend to disseminate our work by preparing scientific publications for international scientific journals and conferences, once our project is complete.

6. Confirmation by the host institution of the successful execution of the STSM

Please see attached.

7. Other comments

I am very grateful to Ass. Professor Georgios Petrou and his team for the kind welcome and to the KEYSTONE COST action members to make this STSM possible.