

SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: IC1302 - KEYSTONE

STSM title: Keyword search on big data

STSM start and end date: 03/07/2017 to 21/07/2017

Grantee name: Nikolche Spasevski

PURPOSE OF THE STSM:

(max.200 words)

The purpose of the STSM was to enable Mr.Nikolche Spasevski to visit Professor Ioannis Doxiadis at the National and Kapodistrian University of Athens in order to discuss his experience and to extend his knowledge in the area of keyword search, more specifically in the area of keyword search in big graphs and team formation in social networks.

The main goal of my visit was to: 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 Computer Systems Department at National and Kapodistrian University of Athens, as well as experts in the field of keyword search, graph theory, Social Networks, and Bid Data.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

(max.500 words)

During the STSM, I spent three weeks at the Computer Systems Department of the National and Kapodistrian University of Athens, under supervision of Professor Ioannis Doxiadis. 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 keyword search, graph theory, Social Networks, and Bid Data, 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 discussion on the problems, challenges and opportunities for improving the exploration of graph-like databases. Keyword search over graph-like databases offers an alternative way to access and use semi-structured data that neither requires mastery of a query language, nor deep knowledge of the database's potentially quite complex schema. More specifically, the work carry out during the STSM revolved around building systems to resolve these two problems: keyword search in big graphs and team formation in social networks.

We agree that we can continue with our collaboration after the finishing of the STSM. 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 Professor Ioannis Doxiadis, we plan to collaborate with the host institution in the near future, in order to achieve the future objectives of our work. 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.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

The main results have been conducted mainly together with my host Professor Ioannis Doxiadis. We managed to achieve all activities attached to the proposed work plan including:

- *Review, identify and analyze the general problems and modern trends in the area of keyword search and big data analysis with special focus on graph-like databases, algorithms, tools and application for this purpose.*
- *Define the semi-structured databases problem as a multiobjective optimization problem and implement the base sequential procedure along with the set of big data sets used for the experiments.*
- *Design a graph-like database procedure for multiobjective optimization based on the island model, with different alternatives for exploring separate search spaces across different subpopulations.*
- *Implement the graph-like database procedure and evaluate the implemented search space strategies with the alternative objective formulations used to determine the quality of the unsupervised classification obtained with the selected features.*

The initial experimental results have shown good 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 big data and social networks, and finalize our research timetable and intended output plans.

FUTURE COLLABORATIONS (if applicable)

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 Professor Ioannis Doxiadis, we plan to collaborate with the host institution in the near future, in order to achieve the future objectives of our work. 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.