



LUNDS
UNIVERSITET

September, 2014

Master Thesis Proposal

Network Analysis of Apache Software Foundation Source Code Repositories (2 students)

Context:

During the last decade, Open Source Software (OSS) has experienced a significant increase in number of users and contributors, especially the ones coming from the software heavy industries, e.g. Google, Ericsson, and HP. Apache Software Foundation hosts over 150 OSS projects, majority of which can be characterized as mature, large scale, and industry grade software products.

The availability of OSS source code and communication archives has enabled researchers to study large-scale software development in more detail, e.g., how participants organize and collaborate, underlying development practices, and evolution of the software architectures. The results of such studies can be applicable within industry setting such as the ongoing trend of open source development practices implementation within a closed setting, also known as innersource. Network theory has been applied to analyze OSS source code repositories and communication networks.

The goal of this master thesis project is to obtain some source code related data, verify and analyze it, calculate network metrics and categorize the obtained results for all OSS projects hosted under the ASF. The created metrics should then serve as a benchmark for other software projects, e.g., proprietary software projects developed within a closed company setting. Looking at how network metrics evolve over time, and relating them to a software company development process changes could be used to monitor and improve the underlying development methodologies.

This project is a mix of design, programming, and data analysis by applying network theory and using Gephi network analysis software.

Student Profile: Excellent Java programming skills and knowledge of relational dbms, and version control systems such as Git and SVN. Basic knowledge of statistics is required, while knowledge of network theory is a plus.

Please contact Alma Orucevic-Alagic (alma@cs.lth.se)