A Visual Framework for Graph and Text Analytics in Email Investigation

Abstract

The aim of this work is to build a framework which can benefit from data analysis techniques to explore and mine important information stored in an email collection archive. The analysis of email data could be accomplished from different perspectives, we mainly focused our approach on two different aspects: social behaviors and the textual content of the emails body. We will present a review on the past techniques and features adopted to handle this type of analysis, and evaluate them in real tools. This background will motivate our choices and proposed approach, and help us build a final visual framework which can analyze and show social graph networks along with other data visualization elements that assist users in understanding and dynamically elaborating the email data uploaded. We will present the architecture and logical structure of the framework, and show the flexibility nature of the system for future integrations and improvements. The functional aspects of our approach will be tested using the 'enron dataset', and by applying real key actors involved in the 'enron case' scandal.