Temporal Visualization of Dynamic Collaboration Graphs of OSS Software Forks

Amir Azarbakht, Carlos Jensen

Oregon State University, School of Electrical Engineering & Computer Science
1148 Kelley Engineering Center, Corvallis OR 97331, USA
{azarbaam, cjensen}@eecs.oregonstate.edu
http://eecs.oregonstate.edu/~azarbaam
http://eecs.oregonstate.edu/people/jensen

Keywords. Temporal Visualization, Temporal Analysis, Social Dynamics, Free and Open Source Software, Social Network Analysis

Abstract

In this work, we studied collaboration network of three open source projects using a combined analysis method of temporal visualization and temporal quantitative analysis. We based our study on two papers by [Robles and Gonzalez-Barahona 2012] and [Hanneman and Klamma 2013], and identified three projects that had forked in the recent past. We mined the collaboration data, formed dynamic collaboration graphs, and measured social network analysis metrics over an 18-month period time window.

We also visualized the dynamic graph (available online) and as stacked area charts over time. The visualizations and the quantitative results showed the differences among the projects in the three forking reasons of personal differences among the developer teams, technical differences (addition of new functionality) and more community-driven development. The personal differences representative project was identifiable, and so was the date it forked, with a month accuracy.

The novelty of the approach was in applying the temporal analysis rather than static analysis, and in the temporal visualization of community structure. We showed that this approach shed light on the structure of these projects and reveal information that cannot be seen otherwise.
Fig. 1: Snapshots from the video visualization of Kamailio project’s collaboration graph (October 2007 - March 2009) in which a core contributor (colored red) moves to the periphery and eventually departs out of the community.

Fig. 2: Kamailio project’s top contributors’ betweenness centralities and network diameter over time between October 2009 to March 2011 in 3-months time windows with 1.5 months overlaps.