

The micro-dynamics of open source software development activity

Paul A. David¹, Francesco Rullani²

1 Stanford University, Economics Bldg 333
Stanford, California, 94305-6072, United States; and All Souls College,
Oxford, OX1, 4AL, United Kingdom.

2 Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà 33,
56127 Pisa, Italy.
pad@stanford.edu, rullani@sssup.it

Abstract. This study aims to isolate and identify the properties of FLOSS development insofar as these can be revealed by examining the ecology of *SF.net*. It characterizes the contrast between the many “lurkers” and a much smaller core of “entrepreneurial” developers who are responsible for launching new projects, and gives an interpretation of the function of platforms such as *SF.net* as sites that people with a propensity to start open source projects can use to recruit “laborers”. It describes the process underpinning the mobility of those who are recruited among the projects that are launched and provides insights on the evolution of developers’ level and mode of involvement in FLOSS production.

1. Research Questions

The FLOSS model has given rise to a self-organizing global ecology of atomistic and collective projects that both share and compete for productive resources as well as for final “users”. This structure is manifested within the microcosm of the SourceForge.net (*SF.net* henceforth), the largest platform for FLOSS development worldwide. By studying a dataset containing information about the population of 222,835 developers who registered themselves on *SF.net* during an early period in the platform’s history (specifically, from September 1, 2000 through October 26, 2001), we have been able to address the following questions.

Is there a stable, distinct typology of actors, e.g. “entrepreneurial” developers who launches many projects, “laborers” who participate as group members of existing projects without launching any projects themselves, and “lurkers”, who simply observe or contribute from outside the projects teams? How and on the basis of which characteristics individuals move over time between lurking, laboring, and launching projects?

2. Analysis

We define the following 7 states and assign each developer to one of these for every 30-days period of his/her “life” in *SF.net*:

0=non member and non founder, inactive (i.e. she/ he did not post any bug report,

Please use the following format when citing this chapter:

David, P.A., and Rullani, F., 2006, in IFIP International Federation for Information Processing, Volume 203, Open Source Systems, eds. Damiani, E., Fitzgerald, B., Scacchi, W., Scotto, M., Succi, G., (Boston: Springer), pp. 339-340

patch or feature request);

1=non member and non founder, active;

2=member of 1 project and non founder of any project (active or inactive);

3=member of more than 1 projects and non founder of any project (either active or inactive);

4=founder of 1 project and member of 1 project (either active or inactive);

5=founder of 1 project and member of more than 1 projects (either active or inactive);

6=founder of more than 1 project and member of more than 1 projects (either active or inactive).

We study the transitions from one state to the others by applying Markov chain theory to describe how developers' involvement changes over time, obtaining estimates of a series of transition probability matrices pertaining to mutually exclusive sub-samples of developers, spanning the most important characteristics provided in the dataset: developers' *registration date* to *SF.net*; developers' state in *pre-analysis* periods; developers' characteristics such as the *skills* they declare to have, the main *language* they declare to speak, and the provision or not of an *email* address to be directly contacted.

3. Results and Limitations

Comparisons among the obtained strata enable us to better understand the nature of the mechanism triggering the launching of new projects, the participation in existing projects (i.e. the laboring activity) and "passive" participation (i.e. lurking activities).

In particular we have found that early-registered users have higher persistence in their foundation activity, and that developers who send "signals" into the community (e.g. disclosing their skills sets or their email addresses) tend to be, and become over time, more active and "entrepreneurial". The level of pre-analysis activity induces a sort of "role exchange" between the developers, where initially active individuals become inactive, and vice versa. Eventually, languages differences also matter. While English speakers follow the whole population dynamics, European, and even more so Asian, languages speakers enter mainly as lurkers, and then move in a greater proportion to more active states.

The main limitation of the study concerns the focus on the SourceForge population, which is only a sample of the universe of FLOSS projects, and one that does not capture phenomena characteristic of the very large projects. These points to the need, and the opportunities to apply the methodology developed here to the study of other sites, such as Savannah, and FreshMeat.