

# Exploring Characteristics of Code Churn

Jos Kraaijeveld and Eric Bouwers

Technische Universiteit Delft, Software Improvement Group

`mail@kaidence.org`, `e.bouwers@sig.eu`

Software is a centerpiece in today's society. Because of that, much effort is spent measuring various aspects of software. This is done using software metrics. Code churn is one of these metrics. Code churn is a metric measuring change volume between two versions of a system, defined as sum of added, modified and deleted lines. We use code churn to gain more insight into the evolution of software systems. With that in mind, we describe four experiments that we conducted on open source as well as proprietary systems [1].

First, we show how code churn can be calculated on different time intervals and the effect this can have on studies. This can differ up to 20% between commit-based and week-based intervals. Secondly, we use code churn and related metrics to automatically determine what the primary focus of a development team was during a period of time. We show how we built such a classifier with a precision of 74%. Thirdly, we attempted to find generalizable patterns in the code churn progression of systems. We did not find such patterns, and we think this is inherent to software evolution. Finally we study the effect of change volume on the surroundings and user base of a system. We show there is a correlation between change volume and the amount of activity on issue trackers and Q&A websites.

## References

- [1] J. M. Kraaijeveld. Exploring Characteristics of Code Churn. Master's thesis, Delft Technical University, The Netherlands, 2013.