Identifying the Intensity of Variability Changes in Software Product Line Evolution

Christian Kröher\textsuperscript{1} Lea Gerling\textsuperscript{2} Klaus Schmid\textsuperscript{3}

Abstract: This extended abstract summarizes the paper Identifying the Intensity of Variability Changes in Software Product Line Evolution [KGS18] published in the proceedings of the SPLC 2018 [BBB+18].

Keywords: Software product line evolution; evolution analysis; variability changes; intensity

A particular focus of Software Product Line (SPL) research is on understanding the evolution of variability information. This type of information realizes the configuration knowledge enabling the customization of artifacts for a specific product of a SPL, like the configuration options in the variability model or explicit references to these options in pre-processor statements. However, this research is typically based on a feature-perspective, which considers changes only in relation to a specific feature and, hence, abstracts from the implementation details of SPL evolution. As a consequence, the frequency with which developers generally change a specific amount of variability information in different types of artifacts independent from its relation to a specific feature is currently unknown. We believe that this also hides potential for improving current SPL analysis and verification approaches towards a better evolution support. In particular, the intensity (the frequency and amount) of changes to a particular type of artifact and information defines how often certain SPL analyses need to be redone in an evolutionary setting. Hence, we raise and answer the following research questions in this paper:

\begin{itemize}
  \item \textbf{RQ1} How often do changes affect variability information in general?
  \item \textbf{RQ2} How often do changes affect variability information in code, build, and variability model artifacts?
  \item \textbf{RQ3} How many code, build, or variability model lines containing variability information are changed on average?
\end{itemize}

\textsuperscript{1}University of Hildesheim, Institute of Computer Science, Universitätsplatz 1, 31141 Hildesheim, Germany kroher@sse.uni-hildesheim.de
\textsuperscript{2}University of Hildesheim, Institute of Computer Science, Universitätsplatz 1, 31141 Hildesheim, Germany gerling@sse.uni-hildesheim.de
\textsuperscript{3}University of Hildesheim, Institute of Computer Science, Universitätsplatz 1, 31141 Hildesheim, Germany schmid@sse.uni-hildesheim.de

doi:10.18420/se2019-31
In order to answer these questions, we first introduce a fine-grained and variability-centric evolution analysis approach, which provides the necessary details about the evolution of code, build, and variability model artifacts in C-preprocessor-based SPLs. Second, we present the results of applying this approach to the Linux kernel. These results cover more than 12 years of active development as documented by the Git repository\(^4\) between 2005 and 2017. Finally, we discuss the results of this application in relation with the research questions, which reveals the following answers:

**RQ1** Changes to variability information occur in about 11% of all commits from which only around 2% change this information exclusively. The majority of the commits (about 80%) exclusively changes information, which is not related to the configuration knowledge.

**RQ2** Most of the changes to variability information affect code artifacts (7%) followed by variability model artifacts (5%) and build artifacts (3%).

**RQ3** For code and build artifacts, only about 2-3% of the commits change 1-10 lines containing variability information, while for the variability model artifacts such changes occur only slightly more frequently (4%). Commits changing more lines at once occur even less frequently.

In summary, these results show that changes to variability information occur surprisingly infrequently and only affect small parts of the analyzed artifact types. On this basis, we further outline opportunities for improving approaches and tools to support SPL evolution. The main benefit of considering fine-grained information about the actual changes to variability information in the design of new approaches for evolutionary analysis is the reduction of analysis effort.

**Acknowledgments:** This work is partially supported by the Evoline project, funded by the DFG (German Research Foundation) under Priority Programme SPP 1593 and by the ITEA3 project REVaMP\(^2\), funded by the BMBF (German Ministry of Research and Education) under grant 01IS16042H. Any opinions expressed herein are solely by the authors and not of the DFG or BMBF.

**References**


\(^4\) https://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git