

Comparing manual and automated feature location in conceptual models: A Controlled experiment

Francisca Pérez, Jorge Echeverría; Raúl Lapeña, Carlos Cetina
Universidad San Jorge. SVIT Research Group, Autovia, A-23 Zaragoza-Huesca Km.299, Zaragoza 50830, Spain
{mfperez, jecheverria, rlapena, ccetina}@usj.es

Palabras Clave: Controlled experiment, Feature location, Model-Driven engineering, Conceptual models

Lugar de publicación: Information and Software Technology. 125 (2020). 106337

Impact factor: 2.726 – Q2 JCR (índice 2019, categoría Computer Science, Software Engineering)

DOI: <https://doi.org/10.1016/j.infsof.2020.106337>

Resumen(Abstract).

Context: Maintenance activities cannot be completed without locating the set of software artifacts that realize a particular feature of a software system. Manual Feature Location (FL) is widely used in industry, but it becomes challenging (time-consuming and error prone) in large software repositories. To reduce manual efforts, automated FL techniques have been proposed. Research efforts in FL tend to make comparisons between automated FL techniques, ignoring manual FL techniques. Moreover, existing research puts the focus on code, neglecting other artifacts such as models.

Objective: This paper aims to compare manual FL against automated FL in models to answer important questions about performance, productivity, and satisfaction of both treatments.

Method: We run an experiment for comparing manual and automated FL on a set of 18 subjects (5 experts and 13 non-experts) in the domain of our industrial partner, BSH, manufacturer of induction hobs for more than 15 years. We measure performance (recall, precision, and F-measure), productivity (ratio between F-measure and spent time), and satisfaction (perceived ease of use, perceived usefulness, and intention to use) of both treatments, and perform statistical tests to assess whether the obtained differences are significant.

Results: Regarding performance, manual FL significantly outperforms automated FL in precision and F-measure (up to 27.79% and 19.05%, respectively), whereas automated FL significantly outperforms manual FL in recall (up to 32.18%). Regarding productivity, manual FL obtains 3.43%/min, which improves automated FL significantly. Finally, there are no significant differences in satisfaction for both treatments.

Conclusions: The findings of our work can be leveraged to advance research to improve the results of manual and automated FL techniques. For instance, automated FL in industry faces issues such as low discrimination capacity. In addition, the obtained satisfaction results have implications for the usage and possible combination of manual, automated, and guided FL techniques.