

## Detection of a dual licensing violation

Miloš Cvetanović, Zaharije Radivojević, Saša Stojanović  
*School of Electrical Engineering*  
*University of Belgrade*  
{cmilos, zaki, stojsasa}@etf.bg.ac.rs

### Abstract

Dual licensing allows source code to be used free of charge for non-commercial purposes while some fee applies otherwise. An important step in a detection of license violation is a time-consuming search whether a code extracted from a final product contains a procedure that originates from a source code for which a license is required. Still, there is no available tool for performing appropriate search, however, some tools capable to detect different types of software clones could be used.

The authors proposed an approach that uses software metrics for assessing similarities between procedures. Depending on a configuration, extracted software metrics are either further exercised with machine learning classifiers or processed with the newly proposed techniques. Moreover, the authors conducted a series of experiments using existing clone detection tools and different configurations of the proposed approach. The experiments are conducted using a dataset based on the open source BusyBox toolset. The evaluation shows that the proposed approach could achieve double recall while retaining adequate precision in comparison to the other available solutions.

**Keywords:** software clones; software metrics; machine learning; licensing; infringement.