Identifying and Generating Missing Tests using Machine Learning on Execution Traces

Frédéric Tamagnan[*1,2]

[1]Franche-Comté Électronique Mécanique, Thermique et Optique - Sciences et Technologies (UMR 6174) (FEMTO-ST) – Université de Technologie de Belfort-Montbeliard, Ecole Nationale Supérieure de Mécanique et des Microtechniques, Centre National de la Recherche Scientifique : UMR6174, Université de Franche-Comté, Université Bourgogne Franche-Comté [COMUE] – 32 avenue de lÓbservatoire 25044 BESANCON CEDEX, France
[2]Orange – Orange Group – France

Résumé

Testing IT systems has become a major bottleneck for many companies. Besides the growing complexity of such systems, shorter release cycles and increasing quality requirements have led to increased verification and validation costs. However, analysis of existing testing procedures reveals that not all artifacts are exploited to tame this cost increase. In particular, customer traces are usually ignored by validation engineers. In this paper, we use machine learning from execution traces (both customer traces and test execution traces) to identify test needs and to generate new tests in the context of web services and API testing. Log files of customer traces are split into smaller traces (user sessions) then encoded into Pandas DataFrames for data analysis and machine learning. Clustering algorithms are used to analyse the customer traces and compare them with existing system tests, and machine learning models are used to generate missing tests in the desired clusters. The tool-set is implemented in an open-source library called Agilkia.

[*]Intervenant