



**Título:** Towards Automatic Labeling of Exception Handling Bugs

**Data:** 25/11/2022

**Horário:** 14h00

**Local:** Sala de Seminários – Bloco 952

**Resumo:**

Exception handling bugs (EHB) are caused by the incorrect use of the exception handling mechanism (EHM). The vast space of potential error conditions and the sophisticated logic of

contemporary systems (e.g., cloud-based and big data-oriented systems) present a large threat to the correct use of the EHM, which may lead to severe consequences (e.g., system downtime, data loss, and security risk). Thus, this kind of bug must be quickly triaged (i.e., identified, prioritized, and assigned to who best fits) and fixed. However, the identification of EHB is not an easy task, once it mostly relies on the bug reporter's knowledge about the system's error-handling strategy to quickly and correctly classify it. This study aims to evaluate to what extent the combination of Natural Language Processing (NLP) and Machine Learning (ML) techniques can be used to automatically and reliably label EHB by using textual information present in report's summary, description, and comments. Thus, we manually inspected and analyze 4,516 bug reports from four components (Core, HDFS, Yarn, and MapReduce) of Apache's Hadoop project and about 20% (943) of them were labeled as EHB. Then, we perform an evaluation using features extracted from textual fields of bug reports using two NLP techniques, Bag of Words and TF-IDF, and measure the performance of five different ML methods (Support Vector Classifier, Multinomial Naive Bayes, Linear Regression, MultiLayer Perceptron, and Random Florest) when using the two different set of features for the task of automatic labeling of EHB. We also evaluate if using Bag of Words and TF-IDF only on keywords related to exception handling extracted from textual fields could improve the ML models' performance. Our results show that the combination of NLP and ML techniques achieved a good performance in the automatic labeling of EHB, with scores of AUC metric varying from 0.61 up to 0.74. Additionally, considering only keywords related to exception handling the ML models' performance was a little bit improved, with scores of AUC metric varying from 0.65 up to 0.76. To the best of our knowledge, this is the first study addressing the problem of automatic labeling of EHB. Our results can be used to build tools to aid maintainers in the EHB triage process. We also observed that EHB may surf in the class imbalanced problem, which calls for further studies to investigate this threat.

**Banca examinadora:**

- Prof. Dr. Lincoln Souza Rocha (MDCC/UFC - Orientador)
- Prof. Dr. João Paulo Pordeus Gomes (UFC)
- Prof. Dr. Paulo Henrique Mendes Maia (UECE)