

# An Empirical Study on Using Class Stability as an Indicator of Class Similarity

**Mohammad Alshayeb**

Information and Computer Science Department  
King Fahd University of Petroleum & Minerals  
Dhahran 31261, Saudi Arabia  
[alshayeb@kfupm.edu.sa](mailto:alshayeb@kfupm.edu.sa)

## **Abstract**

Software maintenance is an important software quality attribute. Many factors affect software maintenance, one of them being code cloning. Code clones are segments of code that are very similar. Software stability tends to measure the unchanged code elements. The objective of this paper is to find whether stability metrics can be used as an indicator of code structural similarity. I perform an empirical study to find the relationship between code similarity and stability at the class level. I also conduct clustering to classify stability and similarity metrics into different related groups. Finally, I perform principal component analysis (PCA) to determine which class stability metrics have the strongest relationship with class similarity. In addition, I built a prediction model to predict class similarity using class stability metrics. The results show that the four investigated stability metrics have a significant relationship with similarity; however, the Class Stability Metric (CSM) has the strongest correlation with code similarity. The clustering results also reveal that classes with high stability tend to have high similarity. In addition, I found that the CSM and Class Instability Metric (CII) can both reveal 74.023% of class similarity. I conclude that stability metrics can be used as a good indicator of class similarity.

**Keywords:** Class Stability; Class Similarity; Software Metrics; Empirical Study