Service Identification Approach Based on K-Means Algorithm

Ouafa Souiou
LISCO Laboratory, Badji Mokhtar-Annaba University, Annaba, Algeria

Nora Bounour
LISCO Laboratory, Badji Mokhtar-Annaba University, Annaba, Algeria

Abstract:

Service identification is a fundamental task in the legacy system migration to Service-Oriented Architecture (SOA) process. The quality of the identified services has a great impact on the success of the migration in terms of modularity, reusability, and maintainability. In the literature, numerous research efforts have focused on identifying services from legacy software by analyzing source code and applying various clustering techniques. Most of them employed hierarchical clustering due to its intuitive structure and the ability of producing nested service candidates. However, hierarchical methods always suffer from scalability and flexibility problems when applied to large software systems. In this paper, we propose a novel automatic approach for service identification based on the K-means clustering algorithm, which offers a more scalable and efficient alternative. Our method draws inspiration from earlier work in that it employs the same fitness function and service quality metrics, such as cohesion and coupling, to guide the clustering process. Through this work, we aim to demonstrate that K-means can be effectively used to identify high-quality services from legacy source code without compromising but rather improving the effectiveness of traditional hierarchical techniques.

Keywords:

Legacy System, Reengineering, SOA, Clustering, K-Means.