Abstract

Code cloning is the procedure where the developers reuse the code fragments implementing the paste option. They may or may not make the modification in the source code. The code thus developed after copying is known as clone. It is the synonym of duplicate. In the year 2002, Ira Baxter coined the term clones as the segments of code that are similar according to some definition of similarity. The similarity can be based on text, syntactic or semantic. Studies have revealed that almost 10-15% of the source code in large software are part of single or more clones[1]. Clones have adverse impact on the software maintenance, thus identification of clones is beneficial. In the past decade many tools have been developed to detect the clones but none was able to correctly identify all types of clones. In this paper the literature survey of all the clone detection techniques has been done. Along with this it also propose an approach which will use a combination of tree and token based approach in order to detect the code clones.

References


27. Tool SimScan http://www.blue-edge.bg/download.html

28. Project Bauhaus http://www.bauhaus-struggart.de


1993.

Index Terms

Computer Science  Software Engineering

Keywords
Clones, textual comparison, LWH approach, token based approach, PDG approach, metric comparison, AST approach.