Abstract

If two fragments of source code are identical or similar to each other, they are called code clones. Code clones introduce difficulties in software maintenance and cause bug propagation. Software clones occur due to several reasons such as code reuse by copying pre-existing fragments, coding style, and repeated computation using duplicated functions with slight changes in variables or data structures used. If a code fragment is edited, it will have to be checked against all related code clones to see if they need to be modified as well. Removal, avoidance or refactoring of cloned code are other important issues in software maintenance. However, several research studies have demonstrated that removal or refactoring of cloned code is sometimes harmful. In this study, code clones, common types of clones, phases of clone detection, the state-of-the-art in code clone detection techniques and tools, and challenges faced by clone detection techniques are discussed.
24. X. Yan, J. Han, and R. Afshar. Clospan: Mining closed sequential patterns in large datasets, 2003.
38. Weiser, Mark. Program slicing. Proceedings of the 5th international conference on 

39. Hummel, Benjamin, et al. Index-based code clone detection: incremental, distributed, 

Detection and Removal of Code Clones in C and Java Programs using Refactoring. In IJJCET, 
pp. 93-95, 2010.


42. Juillerat, Nicolas, and Bat Hirsbrunner. An algorithm for detecting and removing clones 
in java code. Proceedings of the 3rd Workshop on Software Evolution through Transformations: 

43. Lague, Bruno, et al. Assessing the benefits of incorporating function clone detection in a 

44. Yuan, Yang, and Yao Guo. CMCD: Count matrix based code clone detection. Software 
Engineering Conference (APSEC), 2011 18th Asia Pacific. IEEE, 2011.

using multiple detection results. Software Engineering Conference, 2009. APSEC'09. Asia- 


48. Abd-El-Hafiz, Salwa K. A metrics-based data mining approach for software clone 
detection. Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th 


52. Han, Jiawei. Data Mining: Concepts and Techniques. (2006).

53. Burd, Elizabeth, and John Bailey. Evaluating clone detection tools for use during 

54. Rysselberghe, Filip Van, and Serge Demeyer. Evaluating clone detection techniques 
from a refactoring perspective. Proceedings of the 19th IEEE international conference on 

55. Mayrand, Jean, Claude Leblanc, and Ettore M. Merlo. Experiment on the automatic 
detection of function clones in a software system using metrics. Software Maintenance 1996, 

56. Schleimer, Saul, Daniel S. Wilkerson, and Alex Aiken. Winnowing: local algorithms for 
document fingerprinting. Proceedings of the 2003 ACM SIGMOD international conference on 


**Index Terms**

Computer Science  
Software Engineering
Keywords

Software Clone, Code Clone, Duplicated Code Detection, Clone Detection