1. FaCoY - A Code-to-Code Search Engine

Kisub Kim (1); Dongsun Kim (1); Bissyande, T.F. (1); Eunjong Choi (2); Li Li (3); Klein, J. (1); Le Traon, Y. (1)
Source: *2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE). Proceedings*, p 946-57, 2018; **ISBN-13:** 978-1-4503-5638-1; **DOI:** 10.1145/3180155.3180187; **Conference:** 2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE), 27 May-3 June 2018, Gothenburg, Sweden; **Publisher:** IEEE Computer Society, Los Alamitos, CA, USA
Author affiliation: (1) University of Luxembourg, SnT, Luxembourg (2) Nara Institute of Science and Technology, Japan (3) Monash University, Faculty of Information Technology, Clayton, VIC, Australia
Abstract: Code search is an unavoidable activity in software development. Various approaches and techniques have been explored in the literature to support code search tasks. Most of these approaches focus on serving user queries provided as natural language free-form input. However, there exists a wide range of use-case scenarios where a code-to-code approach would be most beneficial. For example, research directions in code transplantation, code diversity, patch recommendation can leverage a code-to-code search engine to find essential ingredients for their techniques. In this paper, we propose FaCoY, a novel approach for statically finding code fragments which may be semantically similar to user input code. FaCoY implements a query alternation strategy: instead of directly matching code query tokens with code in the search space, FaCoY first attempts to identify other tokens which may also be relevant in implementing the functional behavior of the input code. With various experiments, we show that (1) FaCoY is more effective than online code-to-code search engines; (2) FaCoY can detect more semantic code clones (i.e., Type-4) in BigCloneBench than the state-of-the-art; (3) FaCoY, while static, can detect code fragments which are indeed similar with respect to runtime execution behavior; and (4) FaCoY can be useful in code/patch recommendation. (0 refs)
Inspec controlled terms: query processing - search engines - software engineering - software maintenance
Uncontrolled terms: code-to-code search engine - code search tasks - code-to-code approach - code-patch recommendation - FaCoY - semantic code clones - directly matching code query tokens - user input code - statically finding code fragments - code diversity - code transplantation
Classification Code: C6110B Software engineering techniques - C7250N Search engines
IPC Code: G06F9/44 - G06F17/30
Treatment: Bibliography (BIB) - Practical (PRA)
Database: Inspec
Data Provider: Engineering Village