Code clone detection algorithm for software updates and plagiarism analysis

Technology #cu15182

Code clone detection is a central tool for software update technology and code plagiarism detection. Due to the complexity of detection, current methods often utilize simplified algorithms that fail to identify all exact-match and near-match clones. This technology provides a dynamic algorithm that can detect exact- and near-match code clones while optimizing computation time. With this technology, software updates and source code plagiarism detection can be performed with higher efficiency, speed, and accuracy.

Robust exact- and near-match clone detection with minimal computational cost

Computational cost hampers many existing approaches to clone detection. To combat time and computationally intensive algorithms, current approaches utilize simplified data representations for their detection schemes, which fail to fully identify all existing clones. This proposed technology uses the complex but complete graph data structure to represent code. To do so, it employs a novel algorithm to efficiently traverse the data structure for clone detection. The graph structure is more robust, as it records relationship dependencies within code. This technology is able to efficiently compare these dependencies to reference code (link analysis) to determine if clones exist. Additionally, existing methods are only able to determine exact matches. This technology is also able to determine near-match clones by relying on link analysis.

A prototype of the technology has been tested and has been shown to be robust in identifying exact and near match clones.

Lead Inventor:

Simha Sethumadhavan, Ph.D.
Applications:

- Data mining
- Code refactoring for API updates
- Open-source code plagiarism detection

Advantages:

- Exact and near match clone capabilities
- Provides large search space of possible clones
- Optimizes computation time

Patent Information:

Patent Pending (WO/2016/130542)

Tech Ventures Reference: IR CU15182

Related Publications:


Inventors

Lakshminarasimhan Sethumadhavan