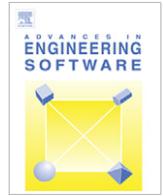




Contents lists available at ScienceDirect

## Advances in Engineering Software

journal homepage: [www.elsevier.com/locate/advengsoft](http://www.elsevier.com/locate/advengsoft)

## Software refactoring at the function level using new Adaptive K-Nearest Neighbor algorithm

Abdulaziz Alkhalid<sup>a</sup>, Mohammad Alshayeb<sup>b,\*</sup>, Sabri Mahmoud<sup>b</sup>

<sup>a</sup> 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

<sup>b</sup> Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

### ARTICLE INFO

#### Article history:

Received 2 April 2010

Received in revised form 27 June 2010

Accepted 5 August 2010

#### Keywords:

Software refactoring

Clustering

Cohesion

Code restructuring

Function level cohesion

Software quality

### ABSTRACT

Improving the quality of software is a vital target of software engineering. Constantly evolving requirements force software developers to enhance, modify, or adapt software. Thus, increasing internal complexity, maintenance effort, and ultimately cost. In trying to balance between the needs to change software, maintain high quality, and keep the maintenance effort and cost low, refactoring comes up as a solution. Refactoring aims to improve a number of quality factors, among which is understandability. Enhancing understandability of ill-structured software decreases the maintenance effort and cost. To improve understandability, designers need to maximize cohesion and minimize coupling, which becomes more difficult to achieve as software evolves and internal complexity increases. In this paper, we propose a new Adaptive K-Nearest Neighbor (A-KNN) algorithm to perform clustering with different attribute weights. The technique is used to assist software developers in refactoring at the function/method level. This is achieved by identifying ill-structured software entities and providing suggestions to improve cohesion of such entities. We also compare the proposed technique with three function-level clustering techniques Single Linkage algorithm (SLINK), Complete Linkage algorithm (CLINK) and Weighted Pair-Group Method using Arithmetic averages (WPGMA). A-KNN showed competitive performance with the other three algorithms and required less computational complexity.

© 2010 Elsevier Ltd. All rights reserved.

\* Corresponding author.

E-mail addresses: [abdulaziz.alkhalid@kaust.edu.sa](mailto:abdulaziz.alkhalid@kaust.edu.sa) (A. Alkhalid), [alshayeb@kfupm.edu.sa](mailto:alshayeb@kfupm.edu.sa) (M. Alshayeb), [smasaad@kfupm.edu.sa](mailto:smasaad@kfupm.edu.sa) (S. Mahmoud).