بسم الله الرحمن الرحيم

MONA MAGHRABY

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

MONA MAGHRABY

شبكة المعلومات الجامعية

التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

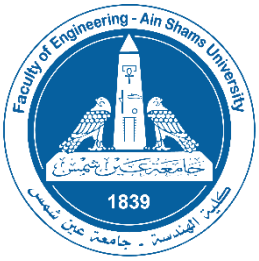
قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

MONA MAGHRABY



AIN SHAMS UNIVERSITY

FACULTY OF ENGINEERING

Computer and Systems Engineering Department

Object-Oriented Design Metrics

A Thesis submitted in partial fulfillment of the requirements of the degree of

Master of Science in Electrical Engineering

(Computer and Systems Engineering)

by

Eng. Farida Ali Mohamed Mahfouz

Bachelor of Science in Electrical Engineering

(Computer and Systems Engineering)

Faculty of Engineering, Ain Shams University, 2016

Supervised By

Prof. Dr. Ashraf Mohamed Mohamed AlFarghaly	**Prof. Dr. Ahmed Hassan Mohamed Yousef**	**Dr. Cherif Ramzi Salama**

Cairo - (2021)

AIN SHAMS UNIVERSITY

FACULTY OF ENGINEERING

Computer and Systems Engineering Department

Object-Oriented Design Metrics

by

Eng. Farida Ali Mohamed Mahfouz

Bachelor of Science in Electrical Engineering

(Computer and Systems Engineering)

Faculty of Engineering, Ain Shams University, 2016

Examiners' Committee

Name and Affiliation	**Signature**
Prof. Dr. Aliaa Abd El-Halim Abd El-Razek College of Computing and Information Technology, Arab Academy for Science, Technology & Maritime Transport (AASTMT)	…….
Prof. Dr. Hoda Korashy Mohamed Ismail Computer Engineering and Systems Faculty of Engineering, Ain Shams University	…….
Prof. Dr. Ashraf Mohamed Mohamed AlFarghaly Computer Engineering and Systems Faculty of Engineering, Ain Shams University	…….
Prof. Dr. Ahmed Hassan Youssef School of Information Technology and Computer Science, Nile University	…….

Date: 13 March 2021

Statement

This thesis is submitted as a partial fulfillment of Master of Science in Electrical Engineering, Faculty of Engineering, Ain shams University.

The author carried out the work included in this thesis, and no part of it has been submitted for a degree or a qualification at any other scientific entity.

Farida Ali Mohamed Mahfouz

Signature

……...…….

Date: 13 March 2021

Researcher Data

Name: Farida Ali Mohamed Mahfouz

Date of birth: 06/04/1993

Place of birth: Cairo, Egypt

Last academic degree: Bachelor of Science

Field of specialization: Electrical Engineering, Computer and Systems Engineering Department

University issued the degree: Ain Shams University

Date of issued degree: July, 2016

Current job: Software Engineer

Abstract

Recently, research studies have been directed to the construction of a universal defect prediction model. Such models are trained using different projects to have enough training data and be generic. One of the main challenges in the construction of a universal model is the different distributions of each software metric among various projects. This study aims to build a universal defect prediction model to predict software defective classes. It also aims to validate the *Object-Oriented Cognitive Complexity* metrics suite (CC metrics) for its association with fault-proneness. Finally, this study aims to compare the prediction performances of each of the CC metrics, the *Chidamber and Kemerer* metrics suite (CK metrics), and a combination of both suites, taking into account the effect of preprocessing techniques applied to them. A neural network model is constructed using three object-oriented metrics sets: the CK metrics, the CC metrics, and a combination of both. Different preprocessing techniques are applied to these metrics sets to overcome the variations in their distributions among various projects. The CK metrics perform well whether a preprocessing technique is applied or not, while the CC metrics' performance is significantly affected by different preprocessing techniques. The CC metrics always outperform in the recall, while the CK metrics usually outperform in the total accuracy, AUC of ROC, precision, F-measure, and MCC. The combination of both the CK and CC metrics exhibits a balance between different performance metrics rather than a superiority in a certain performance metric with a large difference from others. Both quantization and quantization with normalization preprocessing techniques have very close performance. Normalization preprocessing results in the highest recall values using different metrics sets compared to other preprocessing techniques. In conclusion, the construction of a universal model is applicable using different preprocessing techniques and different object-oriented metrics suites. The CC metrics are validated for their association with software fault-proneness. Preprocessing improves the prediction performance when applied to the CC metrics, but it has minimal effect on the prediction performance when applied to the CK metrics.

Thesis Summary

The thesis is organized as follows: Chapter 1 includes an introduction to the research done in the thesis. Chapter 2 explores related work from the literature concerning different aspects, including object-oriented software metrics, software defect prediction, and different data preprocessing techniques used to enhance the defect prediction performance. The software defect prediction section includes both within-project defect prediction and cross-project defect prediction. Chapter 3 reviews the object-oriented software metrics used by the proposed prediction models, the artificial neural network algorithm used to build the proposed model, and the quantiles that are used in the preprocessing of different metrics. Chapter 4 delves into the proposed universal defect prediction model, including details about how datasets have been collected and how metrics are measured. It also includes the explanation of different data preprocessing approaches applied to different metrics sets and details about the construction of the universal defect prediction model. Chapter 5 defines the different metrics used to evaluate the model's performance. Chapter 6 presents and discusses the results of this study. Chapter 7 considers the threats to the validity of this study. Finally, Chapter 8 concludes this research and points to potential future work.

Keywords: Software Defect Prediction, Object-Oriented Metrics, Chidamber and Kemerer Metrics, Cognitive Complexity Metrics, Preprocessing, Universal Model.

Acknowledgment

I would like to thank a lot my supervisors: Prof. Dr. Ashraf M. M. AlFarghaly, Prof. Dr. Ahmed H. M. Yousef, and Dr. Cherif R. Salama for their continuous guidance, help, and support. I am very grateful to have worked with them and their efforts are highly appreciated.

Also, I would like to thank my family, friends, and colleagues for their continuous support, encouragement, and help.

Farida Ali Mohamed
Computer Engineering and Systems
Faculty of Engineering
Ain Shams University
Cairo, Egypt

March 2021

Table of Contents

List of Figures