

Analytical Based Model to Measure Software Engineer's Productivity

Gobikrishnan Thavarajah

(169138G)

Degree of Master of Business Administration in Information Technology

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

June 2018

Analytical Based Model to Measure Software Engineer's Productivity

Gobikrishnan Thavarajah

(169138G)

The dissertation was submitted to the Department of Computer Science and Engineering of the University of Moratuwa in partial fulfilment of the requirement for the Degree of Master of Business Administration in Information Technology.

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

June 2018

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to the University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

.....

Date:.....

Gobikrishnan Thavarajah

The above candidate has carried out research for the Master's thesis under my supervision.

.....

Date

Dr. Amal Shehan Perera

COPYRIGHT STATEMENT

I hereby grant the University of Moratuwa the right to archive and to make available my thesis or dissertation in whole or part in the University Libraries in all forms of media, subject to the provisions of the current copyright act of Sri Lanka. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

.....

Gobikrishnan Thavarajah

ABSTRACT

In the software engineering industry one of the most central business factors is software developer's productivity, the understanding of the term productivity in the context of software development is not clearly defined, however, which cannot be measured cannot be managed, hence, software engineering companies from startup to enterprise are trying their level best to measure software developer's productivity level.

In order to solve this issue, everyone should have an understanding about software engineer's productivity, and also common as well as important factors which could act as an indicator to software developer productivity should be identified and validated. Considering the nature of the problem, a single factor cannot be considered as an indicator of a developer's productivity. Hence a multifactor model should be identified, validated and fine-tuned to produce better accuracy.

As part of this research, a survey among software developers was conducted in order to build a multifactor model which can be used to measure developer's productivity; the model was validated with real software development data and calibrate to producer more accurate result.

ACKNOWLEDGEMENT

It is with immense gratitude that I acknowledge the guidance of my Supervisor, Dr. Shehan Perera. I am indebted to him for his advice, helpful comments and for the suggestions offered to me throughout the research study.

I am especially grateful to Dr. Chandana Gamage and Ms. Jeeva Padmini for sparing valuable time and providing continuous support for the identification of the research problem, formulating the conceptual model, survey data analysis, and reviewing research progress which contributed a lot to carry out the research successfully.

My sincere gratitude is also extended to all the academic and non-academic staff of Department of Computer Science & Engineering, The University of Moratuwa and my colleagues from the MBA in IT degree program for their support in numerous ways.

My thanks are also due to all my friends in 99X Technology, Tiqri Corporation Pvt Ltd, Millennium IT, who helped me in the most important and difficult task of gathering data during the limited period. I would also like to thank all those who responded to my questionnaire.

Gobikrishnan Thavarajah

(169138G)

TABLE OF CONTENTS

DECLARATION.....	I
COPYRIGHT STATEMENT	II
ABSTRACT.....	III
ACKNOWLEDGEMENT.....	IV
TABLE OF CONTENTS	V
LIST OF FIGURES	IX
LIST OF TABLES	X
LIST OF ABBREVIATIONS	XI
1. INTRODUCTION	1
1.1. Background	1
1.1.1. Motivation	3
1.1.2. Research Scope	3
1.2. Problem Statement	4
1.2.1. Research Objectives	5
1.2.2. Research significance.....	5
1.2.3. Outline.....	6
2. LITERATURE REVIEW	7
2.1 The Definition of Productivity	7
2.1.1. The productivity of a developer	9
2.2. Factors Indicating the Quality and Quantity	10
2.2.1. Lines of code	11
2.2.2. Code quality	11
2.2.3. Code complexity	12
2.2.4. Function points.....	13
2.2.5. Defects.....	13
2.2.6. Effort estimation.....	14
2.3. Challenges in Measuring the Productivity	15
2.3.1. Performance measurements of value creation are missing	16
2.3.2. No productivity measurements on an r&d level are found	16
2.3.3. Performance measurement process is missing.....	17
2.4. Agile Software Development Methodology.....	17

2.4.1.	Scrum framework.....	18
2.4.2.	Existing performance metric and KPIs	20
2.5.	Summary	22
3.	RESEARCH METHODOLOGY.....	23
3.1.	Research Problem.....	23
3.1.1.	Research method	24
3.1.2.	Conceptual framework of the research	26
3.2.	Development of Hypotheses.....	28
3.3.	Operationalization	28
3.3.1.	Population and sample selection	30
3.3.2.	The process of data collection.....	32
3.4.	Summary	33
4.	DATA ANALYSIS.....	34
4.1.	Introduction	34
4.2.	Data Gathered from Software Development Activity.....	34
4.3.	Descriptive Statistics for Development Activity Data	35
4.3.1.	Development activity data by age	35
4.3.2.	Sample of software engineers categorized by experience level.....	36
4.3.3.	Sample date of software engineers categorized by gender	37
4.4.	Extraction and transformation of development activity data	37
4.4.1.	Code quantity	37
4.4.2.	Code quality	38
4.4.3.	Code complexity	40
4.4.4.	Work effort.....	41
4.4.5.	Productivity.....	41
4.5.	Testing Hypothesis - Pearson's Correlation Analysis.....	41
4.5.1.	The correlation between code quality and productivity.....	42
4.5.2.	The correlation between code quantity and productivity.....	43
4.5.3.	The correlation between code complexity and productivity	43
4.5.4.	The correlation between actual hours worked and productivity of a software developer.....	44
4.5.5.	The logarithm value of correlation between code quality and productivity.....	45

4.5.6.	The logarithm value of correlation between code quantity and productivity.....	46
4.5.7.	The logarithm value of correlation between code complexity and productivity.....	47
4.5.8.	The logarithm value of correlation between actual hours worked and productivity of a software developer	48
4.6.	Linear Regression Analysis.....	49
4.7.	Reliability of Survey Data.....	50
4.8.	Descriptive Statistics for Survey Demographic Data.....	51
4.8.1.	Sample of software engineers grouped by age.....	51
4.8.2.	Sample of software engineers categorized by gender	52
4.8.3.	Sample of software engineers categorized by experience level.....	53
4.9.	Presentation of Variable Related Sections Information	53
4.9.1.	Quality and software productivity.....	53
4.9.2.	Quantity and software productivity.....	54
4.9.3.	Code complexity for software productivity	55
4.9.4.	Work effort for software productivity.....	56
4.10.	Testing Hypothesis - Pearson's Correlation Analysis	57
4.10.1.	The correlation between code quality and productivity of a software developer	58
4.10.2.	The correlation between code quantity and productivity of a software developer	59
4.10.3.	The correlation between code complexity and productivity of a software developer.....	60
4.10.4.	The correlation between minimal work effort and productivity of a software developer.....	61
4.11.	Summary.....	61
5.	RECOMMENDATIONS AND CONCLUSION.....	63
5.1.	Introduction	63
5.1.1.	Research conclusion one	63
5.1.2.	Research conclusion two.....	63
5.1.3.	Research conclusion three.....	64
5.1.4.	Research conclusion four	64
5.2.	Research Assumptions and Limitations	64
5.3.	Recommendation.....	65

5.4. Suggestion for Further Research	65
REFERENCES	67
APPENDIX A: TITLE.....	71
APPENDIX B: TITLE.....	76

LIST OF FIGURES

Figure 1 Preliminary steps of Research	8
Figure 2 Research method.....	25
Figure 3 Overall IT workforce by job category	31
Figure 4 Process of data collection	33
Figure 5 – Software development team sample grouped by age	35
Figure 6 - A sample of software engineers categorised by experience level	36
Figure 7 - A Sample Date of Software Engineers Categorized by Gender.....	37
Figure 8 - The correlation between Code quality and productivity	42
Figure 9 - The correlation between Code quantity and productivity	43
Figure 10 - Correlation between code complexity and productivity.....	44
Figure 11 -Correlation between minimal work effort and productivity.....	44
Figure 12 -The correlation between Code quality and productivity	45
Figure 13 - The correlation between Code quantity and productivity	46
Figure 14 - Correlation between code complexity and productivity.....	47
Figure 15 - Correlation between minimal work effort and productivity.....	48
Figure 16 - Non-Liner Regression Analysis	49
Figure 17 Model Summary	50
Figure 18 Sample of software engineers grouped by age	51
Figure 19 Sample of software engineers categorized by gender.....	52
Figure 20 Sample of software engineers categorized by experience level	53
Figure 21 Quality and software productivity	54
Figure 22 Quantity and software productivity	55
Figure 23 Code complexity for software productivity.....	56
Figure 24 Work effort for software productivity	57
Figure 25 Survey Questions part 1.....	71
Figure 26 Survey Questions part 2.....	72
Figure 27 Survey Questions part 3.....	73
Figure 28 Survey Questions part 4.....	74
Figure 29 Survey Questions part 5.....	75

LIST OF TABLES

Table 1 Factors in the conceptual framework	26
Table 2 Research hypotheses	28
Table 3 Operationalization.....	28
Table 4 Reliability of surveys data.....	50
Table 5 The correlation between Code quality and productivity.....	58
Table 6 The correlation between Code quantity and productivity.....	59
Table 7 Correlation between code complexity and productivity	60
Table 8 Correlation between minimal work effort and productivity	61
Table 9 Correlation Values	63

LIST OF ABBREVIATIONS

Abbreviation	Description
ICT	: Information and Communication Technology
LOC	: Lines of Code
IT	: Information Technology
SLOC	: Source Lines of Code
CNN	: Cyclometric Complexity Number
R&D	: Research and development
MTBF	: Mean Time Between Failures
CAC	: Cronbach's Alpha Coefficient
SQA	: Software Quality Assurance