Science Arena Publications
Specialty Journal of Electronic and Computer Sciences
Available online at www.sciarena.com
2016, Vol, 2 (4): 21-34

Presenting a Method to Reduce Deviation of Effort Distribution and Increase the Accuracy of Effort Estimation in Software Development Phases

Sayyed Hedayat Tarighi Nezhad,Abozar[1]. Dolatabadi Nezhad[2], , Sayyed Jalalaldin Gharibi Karyak[3]

[1]Faculty of Computer Engineering,Najafabad Branch,Islamic Azad University,Najafabad,Iran
Email: a_doulatabadi@yahoo.com
[2]Department of Computer Engineering Yasooj Branch, Islamic Azad University Yasooj, Iran
Email: heva.tarighi@yahoo.com
[3]Department of Computer Engineering Yasooj Branch, Islamic Azad University Yasooj, Iran
Email:Gharibi.jalal92@gmail.com

Abstract: *Reduction of deviation in effort estimation prevents from time and financial overheads in software development. Studying research conducted on recognition of influential factors on deviation of effort distribution, this study attempts to categorize and analyze the results of the research on ISBSG data. Finally using results of the analyses, the accuracy of effort estimation will be calculated which is obtained by a reduction in the deviation of effort distribution. Having compared research factors, strategies are presented to reduce the deviation in effort distribution. Strategies presented in this study can play a crucial role in reducing time and financial overheads and increase the accuracy of effort estimation in software development management. Research results can be used to present an effort estimation model with higher accuracy in software development.*
Key terms: *effort estimation, effort distribution, software development management, COCOMO II*

INTRODUCTION

One of the important points in software development is accuracy in management that can directly influence the success and quality of a software product [1]. Accuracy in software development management includes various items such as accuracy in planning and project timing. Some of the main challenges that management faces in planning are different types of estimation that should take place at the beginning and during conducting the project by management team. Different models have been presented so far in the field of effort estimation and software cost, each of which has attempted to increase accuracy of effort estimation in a way. The most important one of them is COCOMO II (Constructive Cost Model II) effort estimation model [2]. Despite the attempts in increasing the accuracy of effort estimation in the presented models, the issue of effort estimation and cost in software development is still one of the important challenges in software management. The reason is the uncertain identity and rapid changes in software development methods. One of the items that project management team should carry out after effort estimation is effort distribution

between software development phases. Deviation in effort distribution is another challenge on the way of software development management. Various models have also been presented for effort distribution and MBASE/RUP is one of the most important of those models [3]. Effort distribution in this model is studied in the following figure [4].

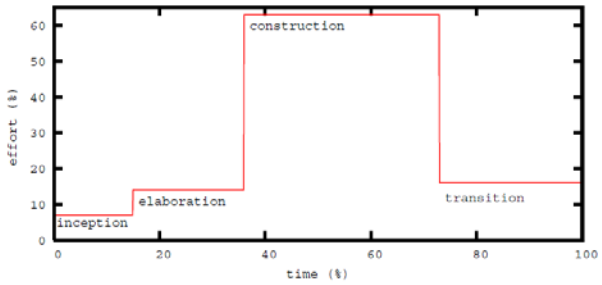

Figure 1: Investigation of effort distribution in MBASE/RUP model

Figure 3: Effort distribution in software development phases based on domain factor in software product

In the studies conducted so far, the reasons for deviation in effort distribution and influential factors in the deviation have been investigated. Results include presentation of strategies to reduce the deviation in effort distribution. None of the previously conducted studies have focused on investigation of the effect of deviation of effort distribution on the accuracy of effort estimation and they have mainly studied the influential factors in deviation of effort distribution. As an example, the effect of software size factor on the deviation of effort distribution is studied in [5] and the sum of its results is presented in the following figure.

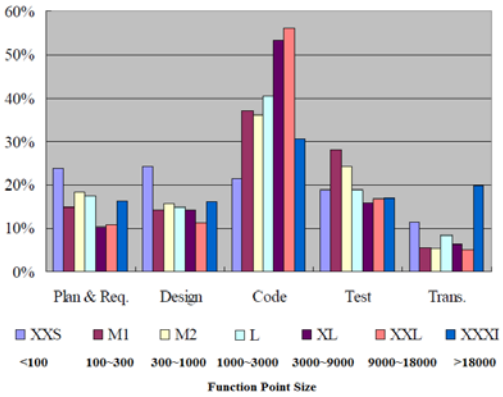

Figure 2: Effort distribution in software development phases based on software size factor

As it can be seen in the above figure, for example, the most effort in software development in XXS (Extra Extra Small) size should be dedicated to the design phase. By investigating the influential factors in deviation of effort distribution on ISBSG (International Software Benchmarking Standards Group) data set, the factors are categorized and analyzed in this study and the of effect of the factors on the accuracy of effort estimation

is calculated. Finally, strategies are presented by the use of analyses to be used by management team in order to reduce the deviation in effort estimation and increase the accuracy of effort distribution.

2. Related Literature

Many studies have been conducted so far on effort estimation and presenting models to increase the accuracy of effort estimation whose results include presentation of different models such as experimental, formal, expert models and so on. One of the most important models that has been presented so far is COCOMO [7] that was first presented in 1981 by Bohem and its last version is COCOMO II that supports object-oriented concepts like the concept of reuse. Fewer studies have been conducted on effort distribution compared to other effort estimation models and the most important one of them presented for effort distribution is presented in the following table [4].

Table 1: The most important effort estimation models in software development

MODEL	THE PHASES (EFFORT DISTRIBUTION PERCENTAGES) COVERED IN DIFFERENT MODELS
COCOMO II [8]	Waterfall distribution scheme: Plan and requirement (7%), preliminary design (17%), detailed design (25%), code (33%), Integration& Test (25), deployment & maintenance (12%), MBASE/RUP distribution scheme :Inception(6%,(Elaboration(24%), Construction(76%), Transition(%12)
COCOMO 81 [7]	Plan and requirement, preliminary design, detailed design (25%), code (33%), Integration & Test (25%)
RUP [9]	Inception(5%), Elaboration(20%), Construction(65%), Transition(10%)

COCOTS [8]	**Assessment, Tailoring, Glue code &Integration. No unified distribution guideline**
SLIM [10]	**Concept Definition, Requirement & Design, Construct & Test, Perfective Maintenance (distribution percentage not available)**
SEER-SEM [11]	**Early specification, design, development, delivery & maintenance (distribution percentage not available)**

Relatively few studies have been conducted on the investigation of the causes of deviation in effort distribution and each study investigates one or several important factors in deviation in effort distribution. The most important studies carried out on recognition and investigation of influential factors in deviation in effort distribution are presented in the following table.

Table 2: Classification of research works on investigation of influential factors in deviation in effort distribution

Researcher	Research year	Research factors	Used data sets
Ye Yang [5]	2008	Software size and type	CSBSG
Thomas Tan [12]	2012	Software size and domain	ISBSG
Qing Wang [13]	2008	Software error	ISBSG

3. Investigation of the Reduction in Deviation in Effort Distribution

Investigating the obtained results from previous studies for the analysis of the of the effect of different factors such as size, development type and domain on the deviation in effort distribution and applying the results of previous studies on evaluation data set in this study, the of reduction in deviation in effort distribution is investigated and extracted in the data set in this section. Results obtained in this section are used in section 4 to investigate the increase in the accuracy of effort estimation which is a result of reduction in the deviation in effort distribution. In the following table, the phases used in this study are classified which are used in the next sections.

Table 3: Classification of research phases for software development

Phase	**Activities**
Requirement(RQ)	Requirement Analysis
Design(AD)	Product Design, Detailed Design
Code(IM)	Code, Unit test, Integration
Test(TE)	System test
Transition(RE)	Installation, Transition, Acceptance Test, User training, Support

3.1. Influential Factors on the Deviation of Effort Distribution

Investigating the influential factors in the deviation of effort distribution in software development, the studies carried out in [5] and [12] are used for size and D-Type factors and the of effect of each one of these factors on the deviation of effect distribution is calculated and presented in the following tables.

Table 4: Investigation of the of effect of size factor on the deviation of effort distribution

Size	**RA**	**AD**	**IM**	**TE**	**RE**
XXS	7.72	9.42	18.88	-2.56	4.30
XS	-1.18	-0.66	-3.29	6.55	-1.42
M1	2.40	0.74	-4.25	2.74	-1.63
M2	1.47	-0.01	-0.08	-2.64	1.27

L	-5.90	-0.55	12.38	-5.69	0.69
XL	-5.32	-3.66	15.74	-4.87	-1.89
XXL	0.63	1.21	-9.69	-4.56	12.87

Table 5: Investigation of the of the effect of D-Type factor on the deviation of effort distribution

D_Type	**RA**	**AD**	**IM**	**TE**	**RE**
Re Development	-3.64	-3.96	8.95	-4.96	3.59
Enhancement	-0.44	-0.12	2.20	-1.07	-0.58
New Development	2.82	1.88	-10.67	5.40	0.56

Investigating the studies carried out in [12] and [14], the effect of domain factor on the deviation of effort distribution is investigated in the following and the sum of the obtained results is presented in the following table.

Table 6: Investigation of the of the effect of domain factor on the deviation of effort distribution

Domain	**RA**	**AD**	**IM**	**TE & RE**
Business	20.98	22.55	24.96	31.51
Command & Control	21.04	22.56	33.73	22.66
Communications	14.95	30.88	28.54	25.62
Control & Display	14.72	34.80	24.39	26.09
Mission Management	15.40	17.78	28.63	38.20
Mission	17.63	12.54	44.32	25.60

Planning				
Sensors Control and Processing	7.78	45.74	22.29	24.19
Simulation	10.71	39.11	30.80	19.38
Spacecraft Bus	33.04	20.66	30	16.30
Weapons Delivery and Control	11.50	17.39	29.82	41.29

All the obtained results in the above tables are finally evaluated on ISBSG data sets and results of evaluation are presented in section 2-3. The data sets used for evaluation are presented in the following table.

Table 7: ISBSG data sets for research evaluation

P#	**Domain**	**Size (KSLOC)**	**D_Type**	**RQ**	**AD**	**IM**	**TE & RE**
A	Command & Control	14.82	New_ DEV	%15	%22	%43	%21
B	Communications	127.5	New_ DEV	%16	%29	%40	%15
C	Simulation	171	Re_DEV	%2	%37	%51.7	%8.1
E	NA	30	NA	%7	%8	%47	%38
F	NA	95	NA	%14	%6	%42	%38
J	NA	50	NA	%15	%9	%48	%28

3.2. Evaluation of Research Data Set

Using the obtained results from the previous section, the reduction in the deviation of effort distribution on ISBSG data set is studied in this section. Results indicating a reduction in the deviation of effort distribution in evaluation projects are obtained by the use of applying and taking the influential factors on the deviation of effort estimation into consideration. The obtained results of effort distribution are compared with COCOMO II effort distribution model. As it can be seen in the following table and diagram, effort distribution with less

deviation than COCOMO II effort estimation model are done by considering factors such as software development size, domain and D-Type. Results obtained from ISBSG data sets are presented in the following table and diagram.

Table 8: Classification of the results of reduction in the deviation of effort distribution in ISBSG data set

Factor	RQ	AD	IM	TE & RE
Domain	%5.16	%1.66	%13.33	%7.06
Size	%8.5	%6.66	%10.83	%18.5
D-Type	%8	%17.8	%16.5	%9

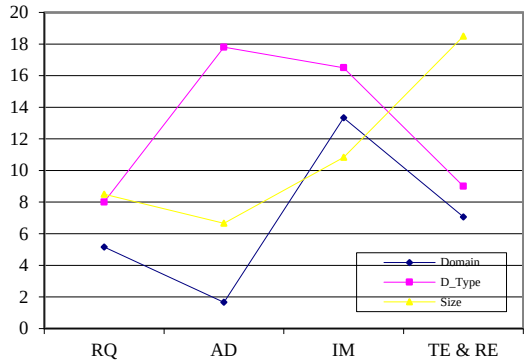

Diagram 1: The mean of reduction in the deviation of effort distribution in ISBSG data set in comparison to COCOMO II effort distribution model

As it can be seen in the above diagram, for example by considering the size factor, the highest of reduction in the deviation of effort distribution is obtained at TE & RE phase. In order to investigate the effect of research factors, the mean of the reduction in the deviation of effort distribution of the used data set is calculated compared to COCOMO II effort distribution model. Results can be seen in the following diagram.

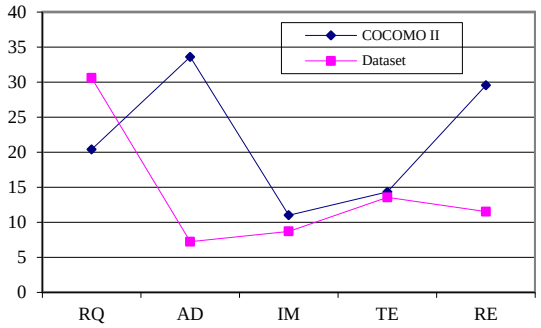

Diagram 2: Investigation of the mean of reduction in the deviation of effort distribution on evaluation data set

As it can be seen in the above diagram, applying the influential factors on the deviation of effort distribution in software development, in average reduces the deviation in effort distribution compared to COCOMO II effort distribution model. Investigating the obtained results in this part, the causes and the effect of deviation of effort distribution on the accuracy of effort estimation are investigated in the next section.

4. Investigation of the Increase in the Accuracy of Effort Estimation

By investigating and analyzing the obtained results from previous section, the increase in the accuracy of effort estimation is studied in this part by the use of the decrease in the deviation of effort distribution. In order to study the increase in the accuracy in the evaluation data set, first the influential factors on the deviation of effort distribution is applied on ISBSG data set and the of reduction in the deviation of evaluation data set is calculated. Results of reduction in the deviation of the evaluation are presented in the following table.

Table 9: Classification of the of deviation of effort distribution in ISBSG data set

Project	RQ	AD	IM	TE & RE
A	%-16	%-7.5	%13.55	%4
B	%1	%12	%13	%11
C	%-13.5	%50	%33	%12.5
E	%0.5	%8	%-23	%20.5
F	%-13	%9.5	%-7.5	%15.5

J	%-12	%-2.5	%2	%19.5

According to the obtained data from the above table and investigation of results obtained from the evaluation, the increased accuracy in effort estimation is extracted. The increased accuracy is calculated in percentage and average among all software development phases. Investigating the increase in the accuracy of effort estimation, according to the obtained results in 2-3, first the reduction in the deviation of effort distribution – as a result of applying the influential factors on the deviation of effort distribution – is extracted. The obtained results are presented in the table and diagram in section 1-3. In the following, the total deviation is extracted which is more and/or less than the estimated effort. The corresponding results can be seen in the previous table. In order to study the of increase in the accuracy of effort estimation, the of effort estimation is calculated in comparison to COCOMO II effort estimation model and the actual effort on three projects of research data sets by applying the influential factors on the deviation of effort distribution. Results of the comparison of the effort estimation and the difference between these values are presented in the following table and diagram.

Table 10: The average of reduction in the deviation of all development phases in comparison with actual and estimated effort distribution

P#	Dataset(PM)	Real Effort(PM)	COCOMO II(PM)
A	962.9	631	1094.3
B	582.6	499	856.7
C	651.1	812.1	586.6

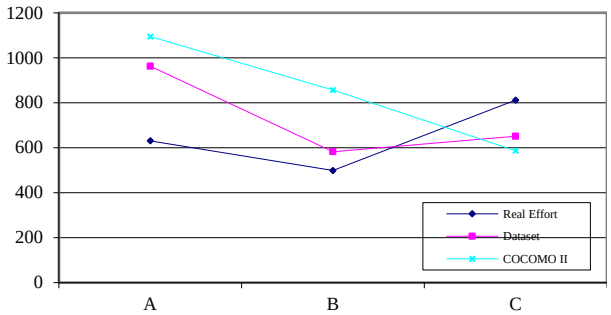

Diagram 3: Effort estimation in research data set in comparison with COCOMO II effort estimation and actual effort

The obtained results in this section and the increase in the accuracy of effort estimation are analyzed in the next part, considering the influential factors on the deviation of effort distribution.

5. Results Analysis

In order to study the increase in the accuracy of effort estimation, the accuracy in effort estimation obtained by applying the influential factors on the deviation of effort distribution is compared with the case where the factors are not applied. The obtained results indicate an increase in the accuracy of effort estimation by applying the influential factors on the deviation of effort distribution. In order to investigate the o accuracy obtained from research evaluation data set, MRE (Magnitude of Relative Error) is used in this section. The following equation is also used to evaluate the level of MRE [15].

MRE = (Actual Effort – Predicted Effort) / Actual Effort

The obtained of MRE metrics from analyzing the evaluation data set is classified in the following table and diagram in comparison with COCOMO II effort estimation model.

Table 11. Investigation of MRE in evaluation data set in comparison to COCOMO II effort estimation model

Project	Dataset MRE	COCOMO II MRE
A	-0.68	-0.73
B	-0.17	-0.71
C	0.20	0.28

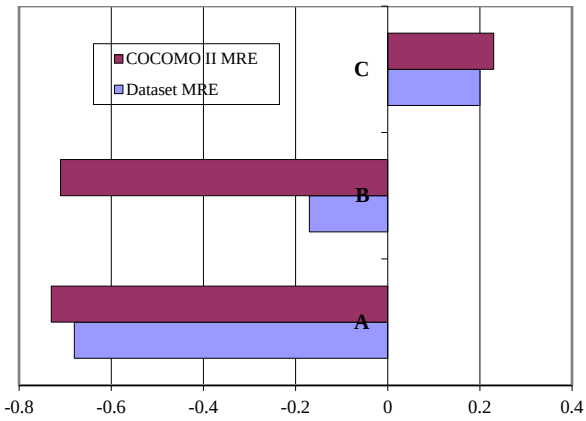

Diagram 4: Investigation of MRE in evaluation data set in comparison to COCOMO II effort estimation model

For the final investigation of the accuracy of effort estimation obtained by the reduction in the deviation of effort distribution (Dataset) in comparison to COCOMO II effort distribution model, MMRE (Mean Magnitude of Relative Error) metrics is investigated in the following and the results are presented in the following table and diagram.

Table 12: Investigation of MMRE in evaluation dataset in comparison to COCOMO II effort estimation

Estimation model	**MMRE**
Dataset	-0.22
COCOMO II	-0.38

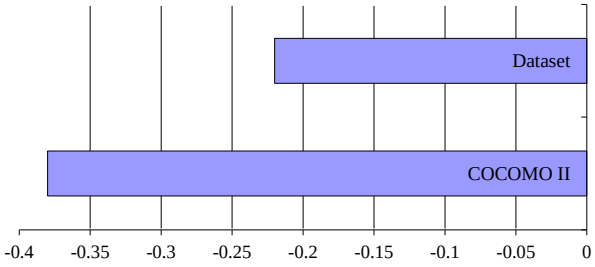

Diagram 5: Investigation of MMRE in evaluation dataset in comparison to COCOMO II effort estimation

The obtained MMRE by applying the influential factors on deviation of effort distribution (DATASET) is less than that of COCOMO II effort estimation and indicates an increase in the accuracy of effort estimation by the use of reduction in the deviation of effort distribution. Based on the analysis results in this section, the

following strategies can be presented for effort distribute and estimation in software development management.

1) If the project management team attempts to distribute effort in software development phases by having an eye on size, domain and D-Type factors, it can lead to a reduction in the deviation of effort distribution and prevent from time and financial overheads in projects.

2) Using effort distribution models like COCOMO II and having an eye on the influential factors on the deviation of effort distribution put forward in this study can finally lead to an increase in the accuracy of effort estimation in software development and an important step in optimized management of software project.

1. Conclusion and Suggestions

In this research work, the effect of the influential factors on the deviation of effort distribution was studied by investigating the influential factors on the deviation of effort distribution in software development. By studying the obtained results from the deviation of effort distribution on evaluation data set, the increase in t accuracy of effort estimation by the reduction in the deviation of effort distribution was analyzed. ISBSG data set was used to evaluate and analyze the results. Evaluation results in the deviation of effort distribution indicate a reduction in the deviation as a result of applying the influential factors on the deviation of effort distribution. The reduction in the deviation of effort distribution can be an important step in preventing time and financial overheads in software development management. Evaluations show an increase in the accuracy of effort estimation by the reduction in the deviation of effort distribution in software development which can have a considerable effect on optimized management of software development. One of the most important research suggestions is investigation of parametric models to apply the influential factors on the deviation of effect distribution. Investigation of other influential factors on the deviation of effort distribution such as commercial environment and volatility of requirements are considered other suggestions for further studies.

References

[1] Menzies, T., Chen, Z., Hihn, J., Lum, K., "Selecting Best Practices For Effort Estimationm", *IEEE Transactions on Software Engineering*, Vol. 32, No. 11, 2006.
[2] Schach, S.R., *Object Oriented &Classical Sofiware Engineering*, Seventh Edition, New York, Mc Graw Hill, 2007.
[3] Boehm, B., Brown, A.W., Fakharzadeh, C., "MBASE/RUP Phase And Activity Distributions (Presentation at COCOMO)", *International Forum 14 Technical Report, 90089, University of Southern California, Los Angeles*, CA,1999.
[4] Heijstek, W., Chaudron, M.R., "Effort Distribution in Model-Based Development", *ACM, 978-1-59593-971*, Kaiserslautern, Germany, 2008.
[5] Yang, Y., Mei, H., Mingshu, L., Wang, Q., Boehm, B.W., "Phase Distribution of Software Development Effort", *ACM, 978-1-59593-971*, 2008.
[6] Hill, P., *Practical Project Estimation*, 2nd Edition. ISBSG, 2005.
[7] Boehm, B., *Software Engineering Economics*, Englewood Cliffs, Prentice-Hall, 1981.

[8] Boehm, B., Abts, C., Brown, A., Chulani, C., Clark, Horowitz, K., *Software Cost Estimation With COCOMO II*, Englewood Cliffs, Prentice-Hall, 2000.

[9] Kruchten, P., "A Brief History of The Rup "Hump Chart", *Technical report, University of British Columbia*, 2003.

[10] Putnam, L., Myers, W., *"Measures for Excellence Yourdon Press Computing Series",* 1992.

[11] SEER-SEM, *http://www.galorath.com/index.php*.

[12] Tan, T., "Domain-Based Effort Distribution Model For Software Cost Estimation", *Phd Thesis, University of Southern California***,** 2012.

[13] Wang**, Q.,** Gou, L., Jiang, N., Che, M., "Estimating Fixing Effort and Schedule Based on Defect Injection Distribution"**,** *Software Process Improvement And Practice, Wiley Inter Science***,** 13: 35–50**,** 2008.

[14] Tan, T., Clark, B., "Technical Report of a New Taxonomy for Software Application Domains and Operating Environments**"**, *USC, CSSE Technical Reports*, 2011.

[15] Cherng, J.,Lin, Y., Tzeng, H., Wang, Y., " Using Computing Intelligence Techniques to Estimate Software Effort", International Journal of Software Engineering & Applications (IJSEA), Vol.4, No.1, 2013.