International Journal of Enhanced Research in Management & Computer Applications, ISSN: 2319-7471
Vol. 4 Issue 5, May-2015, pp: (41-45), Impact Factor: 1.296, Available online at: www.erpublications.com

# Impact of Correlation Analysis in Object-Oriented Technology

Dr. Brij Mohan Goel

Assistant Professor, Vaish College of Engineering, Rohtak

**ABSTRACT**

**Correlation between the classes or within the classes shows the complexity of the design. For one smaller problem, there may be more than one software design but who will be the best; depends on the complexity level of software de-sign. Therefore, correlation which shows the interlinking of classes and strength of classes; control the complexity of the design. The best software object oriented design is based upon the low coupling and high cohesion level. In the present work, a real case study of three Projects InterCafe1, TermoProjekt1 and Zuzel1 demonstrated through the Correlations between the CK Metrics and results are demonstrated in the form of tables.**

**Keywords: Object-Oriented, Correlation, CK Metrics, Coupling, Cohesion.**

## 1. INTRODUCTION

The term coupling is used to measure the relative inter-dependency between various classes as one class has the link with another class. While on the other hand cohesion is defined as the strength of the attributes inside the class which means how the attributes are linked inside the class. Coupling is always correlated with cohesion in such a way as if coupling is high then cohesion is low and vice versa. One can say that a class is highly coupled or many dependent with other classes, if there are many connections and loosely coupled or some dependent with other classes if there is a less connections. The coupling is decided at the designing phase of the system, it de-pends on the interface complexity of the classes. There-fore, the coupling is a degree at which a class is connected with other classes in the system.

Let us now describe the cohesive class which can per-form a single task within the software procedure. It re-quires little interaction with other procedures that are used in other parts of a program. Cohesion gives the strength to the bond between attributes of a class and it is a concept through which capture the intra-module with cohesion. Therefore, cohesion is used to determine how closely or tightly bound the internal attributes of a class to one another. Cohesion gives an idea to the designer about whether the different attributes of a class belong together in the same class. Thus, the coupling and cohesion are related with each other.

## 2. STATISTICS FOR DESCRIPTIVE ANALYSIS OF THE DATA SET

The statistics for the descriptive analysis of the data set is reported using SPSS Package. The following statistics are collected for each of the given metric for three projects namely InterCafe1, TermoProjekt1 and Zuzel1:

- **N**- It shows the number of classes in each project.
- **Min.**- It gives the minimum value of the metric in the project.
- **Max.**- It gives the maximum value of the metric in the project.
- **Mean**- It gives the average of the numbers of a set. Here, it gives the average value of the metric in the project.
- **Standard Deviation** – It is measure of how the numbers are spread out. It is the measure of the dispersion of a data set from its mean value. The deviation is high if the data points are far away from the mean.
- **25%** - It is called the Lower Quartile. It is the point between the lowest 25% of values and the highest 75% of values. It is also known as the **25th percentile**.
- **Median** - It is called the Second Quartile. It gives the middle of the data set. It is also known as the $50^{th}$ percentile.
- **75%** - It is called the Upper Quartile. It is the point between the lowest 75% and highest 25% of values. It is also known as $75^{th}$ Percentile.

The three Projects (NASA data set) InterCafe1, TermoProjekt1 and Zuzel1 are used to predict the fault-proneness of the classes. All these projects are able to predict the fault proneness of classes using six CK metrics.
The following Tables 1, 2 and 3 gives the descriptive analysis of the three projects selected for the analysis.

**Table 1: Descriptive Statistics of the Classes in InterCafe1**

| Metric | N | Min. | Max. | Mean | Std. dev. | 25% | Median | 75% |
|---|---|---|---|---|---|---|---|---|
| WMC | 27 | 2 | 101 | 19.22 | 21.36 | 5 | 10 | 28 |
| RFC | 27 | 4 | 231 | 44.59 | 48.56 | 10 | 30 | 52 |
| CBO | 27 | 1 | 25 | 6.96 | 5.61 | 2 | 6 | 9 |
| LCOM | 27 | 0 | 648 | 96.67 | 164.12 | 1 | 28 | 94 |
| DIT | 27 | 1 | 6 | 2.59 | 2.28 | 1 | 1 | 6 |
| NOC | 27 | 0 | 4 | 0.15 | 0.77 | 0 | 0 | 0 |

**Table 2: Descriptive Statistics of the Classes in Termo Projekt 1**

| Metric | N | Min. | Max. | Mean | Std. dev. | 25% | Median | 75% |
|---|---|---|---|---|---|---|---|---|
| WMC | 42 | 1 | 32 | 4.83 | 5.15 | 2 | 4 | 6 |
| RFC | 42 | 1 | 167 | 20.12 | 29.23 | 3 | 9.5 | 26 |
| CBO | 42 | 0 | 33 | 8.19 | 5.79 | 4 | 7 | 10 |
| LCOM | 42 | 0 | 350 | 13.36 | 54.6 | 0 | 1 | 7 |
| DIT | 42 | 1 | 6 | 2.24 | 1.86 | 1 | 1 | 3 |
| NOC | 42 | 0 | 4 | 0.1 | 0.61 | 0 | 0 | 0 |

**Table 3: Descriptive Statistics of the Classes in Zuzel1**

| Metric | N | Min. | Max. | Mean | Std. dev. | 25% | Median | 75% |
|---|---|---|---|---|---|---|---|---|
| WMC | 29 | 1 | 47 | 13.93 | 10.89 | 5.5 | 11 | 19.5 |
| RFC | 29 | 2 | 95 | 35 | 27.32 | 13 | 28 | 52.5 |
| CBO | 29 | 0 | 42 | 7.48 | 8.44 | 2 | 6 | 10.5 |
| LCOM | 29 | 0 | 707 | 85.1 | 144.48 | 0 | 25 | 107.5 |
| DIT | 29 | 1 | 6 | 2.97 | 2.44 | 1 | 1 | 6 |
| NOC | 29 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## 3. EXPERIMENTAL RESULTS

The results are obtained experimentally by the authors. The results are obtained by investigating the CK metrics on the InterCafe1, TermoProjekt1 and Zuzel1 data set. This data set is preprocessed and is available on NASA repository.
From Table 1, 2 and 3 we have noted that the value of median is low for the metrics DIT and NOC. This shows inheritance is not used much in the considered projects. The value of mean and median is high for the metric LCOM, this indicates that the cohesion of the classes is comparatively low in InterCafe1 (See Table 1) and Zuzel1 (see Table 3) project. The value of mean and median is high for RFC for TermoProjekt1 (see Table 2) project, this indicates that the classes are complex and need more time and effort for testing and maintenance as compared to the small classes. The big differences among the lower 25th percentile, the median, and the 75th percentile are present for all the metrics except the metric NOC. This shows that there exist strong variations across classes. The value of mean and median is slightly high for CBO (not much high), this means the coupling between the classes exists.

**Table 4: Correlations between the Metrics in InterCafe1**

| | | WMC | RFC | CBO | LCOM | DIT | NOC |
|---|---|---|---|---|---|---|---|
| WMC | Pearson Correlation | 1 | .820** | .823** | .766** | .388* | -.142 |
| | Sig. (2-tailed) | | .000 | .000 | .000 | .045 | .478 |
| | N | 27 | 27 | 27 | 27 | 27 | 27 |
| RFC | Pearson Correlation | | 1 | .790** | .780** | .642** | -.151 |
| | Sig. (2-tailed) | | | .000 | .000 | .000 | .453 |
| | N | | 27 | 27 | 27 | 27 | 27 |
| CBO | Pearson Correlation | | | 1 | .623** | .276 | -.177 |
| | Sig. (2-tailed) | | | | .001 | .163 | .377 |
| | N | | | 27 | 27 | 27 | 27 |
| LCOM | Pearson Correlation | | | | 1 | .573** | -.116 |
| | Sig. (2-tailed) | | | | | .002 | .563 |
| | N | | | | 27 | 27 | 27 |
| DIT | Pearson Correlation | | | | | 1 | -.140 |
| | Sig. (2-tailed) | | | | | | .486 |
| | N | | | | | 27 | 27 |
| NOC | Pearson Correlation | | | | | | 1 |
| | Sig. (2-tailed) | | | | | | |
| | N | | | | | | 27 |

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Assuming a reasonably sized data set, According to Hopkins, a correlation value of 0.9–1.0 is almost perfect, 0.7–0.9 very large, 0.5–0.7 large, 0.3–0.5 moderate, 0.1–0.3 minor and less than 0.1 trivial.

Table 4 gives the outcome of linear Pearson's correlations analysis. From Table 4, it is noted that 4 of the 6 metrics are related to each other. It is further concluded that all the six metrics are not completely independent. It is also noted that all the six CK metrics except NOC are considerably related to each other. Further, it is noted that particularly, WMC has a very large correlation with RFC, CBO & LCOM metrics. RFC has a very large correlation with metrics CBO & LCOM. DIT has a moderate correlation with other metrics.

**Table 5: Correlations between the Metrics in TermoProjekt1**

| | | WMC | RFC | CBO | LCOM | DIT | NOC |
|---|---|---|---|---|---|---|---|
| WMC | Pearson Correlation | 1 | .923** | .841** | .893** | .496** | .005 |
| | Sig. (2-tailed) | | .000 | .000 | .000 | .001 | .974 |
| | N | 42 | 42 | 42 | 42 | 42 | 42 |
| RFC | Pearson Correlation | | 1 | .677** | .871** | .666** | -.049 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Sig. (2-tailed) | | | .000 | .000 | .000 | .756 |
| | N | | 42 | 42 | 42 | 42 | 42 |
| CBO | Pearson Correlation | | | 1 | .688** | .238 | .049 |
| | Sig. (2-tailed) | | | | .000 | .129 | .756 |
| | N | | | 42 | 42 | 42 | 42 |
| LCOM | Pearson Correlation | | | | 1 | .425** | -.039 |
| | Sig. (2-tailed) | | | | | .005 | .808 |
| | N | | | | 42 | 42 | 42 |
| DIT | Pearson Correlation | | | | | 1 | -.105 |
| | Sig. (2-tailed) | | | | | | .507 |
| | N | | | | | 42 | 42 |
| NOC | Pearson Correlation | | | | | | 1 |
| | Sig. (2-tailed) | | | | | | |
| | N | | | | | | 42 |

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5 gives the outcome of linear Pearson's correlations analysis. From Table 5, it is noted that 4 of the 6 metrics are related to each other. It is further concluded that all the six metrics are not completely independent. It is also noted that all the six CK metrics except metric NOC are considerably related to each other. Further, it is noted that particularly, metric WMC has almost perfect correlation with metric RFC and very large correlation with metrics CBO and LCOM.

**Table 6: Correlations between the Metrics in Zuzel1**

| | | WMC | RFC | CBO | LCOM | DIT | NOC |
|---|---|---|---|---|---|---|---|
| WMC | Pearson Correlation | 1 | .881** | .448* | .908** | .537** | .[a] |
| | Sig. (2-tailed) | | .000 | .015 | .000 | .003 | . |
| | N | 29 | 29 | 29 | 29 | 29 | 29 |
| RFC | Pearson Correlation | | 1 | .351 | .770** | .778** | .[a] |
| | Sig. (2-tailed) | | | .062 | .000 | .000 | . |
| | N | | 29 | 29 | 29 | 29 | 29 |
| CBO | Pearson Correlation | | | 1 | .269 | .023 | .[a] |
| | Sig. (2-tailed) | | | | .158 | .904 | . |
| | N | | | 29 | 29 | 29 | 29 |
| LCOM | Pearson Correlation | | | | 1 | .470* | .[a] |
| | Sig. (2-tailed) | | | | | .010 | . |
| | N | | | | 29 | 29 | 29 |
| DIT | Pearson Correlation | | | | | 1 | .[a] |
| | Sig. (2-tailed) | | | | | | . |
| | N | | | | | 29 | 29 |
| NOC | Pearson Correlation | | | | | | .[a] |
| | Sig. (2-tailed) | | | | | | |
| | N | | | | | | 29 |

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Table 6 gives the outcome of linear Pearson's correlations analysis. From Table 6, it is noted that 4 of the 6 metrics are related to each other. It is further concluded that all the six metrics are not completely independent. It is also noted that all the six CK metrics except metric NOC are considerably related to each other. Further, it is noted that particularly, WMC hasalmost perfect correlation with LCOM. WMC has very large correlation with RFC. WMC has moderate correlation with CBO.

## CONCLUSION

The outcome of linear Pearson's correlations analysis of CK metrics in three Projects InterCafe1, TermoProjekt1 and Zuzel1 are that all the six CK metrics except NOC are considerably related to each other. DIT has a moderate correlation with other metrics. Inheritance is not used much in the considered projects. The coupling and cohesion between the classes exists.

## REFERENCES

[1] T. M. Meyers and D. Binkley, "An Empirical Study of Slice-Based Cohesion and Coupling Metrics," *ACM Trans- actions on Software Engineering and Methodology*, Vol. 17, No. 1, 2007, Article No. 2.

[2] Y. J. Jeong, H. S. Chae and C. K. Chang, "Semantics Based Cohesion and Coupling Metrics for Evaluating Understandability of State Diagrams," *IEEE* 35*th Annual Computer Software and Applications Conference*, IEEE Computer Society, Washington, 18-22 July 2011.

[3] F. Ensan and W. Du, "A Metric Suite for Evaluating Co-hesion and Coupling in Modular Ontologies," *Proceed-ings of the* 2010 *Conference on Modular Ontologies*: *Pro-ceedings of the* 4*th International Workshop*, IOS Press, Amsterdam, 11 May 2010.

[4] B. Ujhazi, R. Ferenc, D. Poshyvanyk and T. Gyimothy, "New Conceptual Coupling and Cohesion Metrics for Object-Oriented Systems," *Proceedings of the* 2010 10*th IEEE Working Conference on Source Code Analysis and Manipulation*, Timişoara, 12-13 September 2010. doi:10.1109/SCAM.2010.14

[5] S. Husein and A. Oxley, "A Coupling and Cohesion Met-rics Suite for Object-Oriented Software," *Proceedings of the* 2009 *International Conference on Computer Tech-nology and Development*, Vol. 1, IEEE Computer Society, Washington, 2009. doi:10.1109/ICCTD.2009.209

[6] Chowdhury and M. Zulkernine, "Can Complexity, Cou-pling, and Cohesion Metrics Be Used as Early Indicators of Vulnerabilities?" *Proceedings of the* 2010 *ACM Sym-posium on Applied Computing*, ACM, New York, 2010, pp. 1963-1969.

[7] Chowdhury and M. Zulkernine, "Using Complexity, Coupling, and Cohesion Metrics as Early Indicators of Vulnerabilities," *Journal of Systems Architecture*, Vol. 57, No. 3, 2011, pp. 294-313. doi:10.1016/j.sysarc.2010.06.003

[8] S. Oh, H. Y. Yeom and J. Ahn, "Cohesion and Coupling Metrics for Ontology Modules," *Information Technology and Management*, Vol. 12, No. 2, 2011, pp. 81-96. doi:10.1007/s10799-011-0094-5

[9] P. Gandhi and P. K. Bhatia, "Optimization of Object- Oriented Design Using Coupling Metrics," *International Journal of Computer Applications*, Vol. 27, No. 10, 2011, pp. 41-44.

[10] H. Li, Y. He, Z. H. Jiang, Y. C. Huang and Q. Huang, "High-Cohesion and Low-Coupling Integrative Joint for Space Manipulator," *Advanced Intelligent Mechatronics*, *IEEE/ASME International Conference*, Singapore City, 14-17 July 2009, pp. 1463-1467.