



Cophenetic Correlation of Fuzzy Quality Software Defect Analysis

Seetharam.KDept of EEE, Christ University,
Bangalore, India**Basavaraj Chunchure**Dept of CE, SPCOE
Otur, India**Ravindra V. Asundi**Dept of ECE, VIT University
Bangalore, India

Abstract-There are several parameters directly or indirectly quality related to software defects. The value OF these parameters are depending on the size of the software applications. Here 24 variables are Identified all related to quality related to software defects . Defects per function point, Defect Potential, Defect removal efficiency, Defects removed, Defects delivered, Cost per defect prerelease, Cost per defect post releases, Development schedule , Development staffing, Development effort, Development costs, Function points per staff month, LOC per staff month, Maintenance staff, Maintenance effort, Maintenance costs, Total effort, Total cost, Total cost per staff member, Total cost per function point, Total cost per LOC, Average cost per defect, size of c program and function points. The paper discusses qualitative /quantitative measurements of software using cluster analysis. Four different cases are carried out. First analysis as with function point as predominant factor , second analysis with Defects per function point, third analysis with Development schedule and fourth analysis with Development effort [8],[12],[13].

Keywords- Software Quality, cluster, Fuzzy logic, Size, Function point

I. INTRODUCTION

The applications are classified into four groups based on fuzzy logic[14] using the size of function points as described here.[1][5].

Very small applications of 100 function points. Usually have low defect potentials and fairly high defect removal efficiency levels. This is because such small applications can be developed by a single person[9]. There are no interface problems between features developed by different individuals or different teams.

For small applications of 1000 function points, quality becomes very important. At this stage teams are small and methods such as agile development tend to be dominant, other than for systems embedded software. Here team software process (TSP) and the rational unified process (RUP) are more common. In the range of 1000 function points numerous methods are fairly effective.[6]

The quality for large applications when it reaches 10,000 function points they are very significant systems that require close attention to quality control, change control and corporate governance . As application size increases defect potentials increase rapidly and defect removal efficiency level decline with quality control Agile methods are difficult to apply and results are not always good. Cost of quality factor increases as the size increases with 10,000 function points.

The values of quality for very large applications of 100,000 function points range are among the most costly endeavors of modern business. These are also hazardous as many of them fail, and almost all of them exceed their budgets and planned schedules. They are more expensive and troublesome.

II. ANALYSIS

Here four groups are considered [14]. where first group of 100 function points, second group of 1000 function points, third group 10,000 function points, fourth group 100,1000 function points of 24 factors are considered . There are no interface problems between features developed separately for different individuals related to quality values of defects related variables. Each group has five quality defect related factor values are taken. The predictor variables are[3][2][4]:

A. First group with 100 Function points.

- Function point,
- C programming statements
- Defect Potential
- Defects removed
- Defects delivered

B. Second group with 1000 function points

- Defects per function point
- Cost per defect prerelease
- Cost per defect post releases
- Function points per staff month
- LOC per staff month

C. Third group with 10,000 function points of

- Development schedule
- Development staffing
- Maintenance staff
- Maintenance effort
- Total cost

D. Fourth group with 100,000 function points of

- Development effort
- Development, Maintenance costs
- Total effort
- Total cost per staff member
- Total cost per function point

Four different levels, one for each function points group are considered. Five different pair wise distances between observations are considered (Y11, Y21, Y31, Y41, y51). Function points are taken as primary key in every group. Euclidean distance (Euclidean), standard Euclidean distance (Seuclidean), city block distance (city block), minkowski distance (minkowski) and cosine distance. Five different linkage to create a hierarchical cluster tree using five algorithm chosen (Z11, Z21, Z31, Z41, Z51)[7].

‘Single’ nearest distance

‘Complete’, furthest distance

‘Average’, group average distance

‘Weighted’, average distance and

‘Ward’, inner squared distance .

Cophenetic correlation coefficient is used. (C11-...C55) for measuring 25 hierarchical cluster cophenetic correlation coefficient is given in the table.

COPHENET MATRIX

| | | | | | |
|-----|---------|---------|---------|---------|---------|
| | Y11 | Y21 | Y31 | Y41 | Y51 |
| Z11 | 1 | -0.3594 | -0.3172 | 0.6733 | -0.4066 |
| | | | | | |
| | Y21 | Y22 | Y32 | Y42 | Y52 |
| Z21 | 0.3381 | -0.0893 | -0.1377 | 0.1247 | -0.2014 |
| | | | | | |
| | Y31 | Y32 | Y33 | Y43 | Y53 |
| Z31 | -0.2863 | -0.1259 | 0.9938 | -0.3529 | -0.1222 |
| | | | | | |
| | Y41 | Y42 | Y43 | Y44 | Y45 |
| Z41 | 0.6707 | 0.0879 | -0.3593 | 0.9889 | -0.3053 |
| | | | | | |
| | Y51 | Y52 | Y53 | Y54 | Y55 |
| Z51 | 0.2707 | 0.2879 | -0.1352 | 0.4889 | -0.3053 |

III. CONCLUSION

Data is grouped into 4 values applying fuzzy for function points.

Four different group cases are analyzed in detail.

Five different pair wise distances between data.

100,1000,10 000 and 100 000 Function points are taken as primary key.

Each group has five elements in four different levels.

Hierarchical cluster tree is used by five linkage algorithm.

25 Cophenetic correlation coefficient value is calculated.

REFERENCES

- [1] Beck, kent Test driven Development , Boston MA :Addison Wesley 2002
- [2] Chelf, Ben and Raoul Jetley diagnosing medical Device software Defects using static Analysis, San Francisco ,CA : Coverity Technical report, 2008.
- [3] Chess Brain and Jacob west, Secure programming with static analysis Boston, MA: Addison Wesley 2007.
- [4] Cohen, Lou, quality Function deployment –How to make QFD work for you, Upper saddle revering : prentice Hall 1995.
- [5] Crosby, Philip B Quality is Free, New York, NY: New American library, Mentor Books, 1999.
- [6] Everett, Gerald D and Raymond McLeod. Software Testing .Hoboken, NJ: John Wiley& Sons, 2007.

- [7] International organization for standards.ISO 9000/ISO 14000.[http: //www. iso.org /isoen/iso9000-14000/index.html](http://www.iso.org/iso/en/iso9000-14000/index.html).
- [8] Jones capers. Software quality –Analysis and Guidelines for success. Boston, MA: International Thomson Computer Press 1997.
- [9] Kan, Stephen H .Metrics and models in software quality engineering , Second edition .Boston, MA: Addison Wesley Longman, 2003.
- [10] Land, Susan K, Douglas B. Smith, John Z. Walz . Practical Support for Lean Six Sigma Software Process Definition : Using IEEE Software Engineering Standards . Los Almitos, CA; Wiley –IEEE Computer Society Press, 2008.
- [11] Nandyal. Raghav .”Making Sense of Software Quality Assurance”, New Delhi : Tata McGraw Hill Publishing , 2007.
- [12] Capers Jones “Applied Software Measurement”, Tata McGraw Hill Edition 2008.
- [13] Capers Jones “Software Engineering Best Practices”, Tata McGraw Hill Edition 2010.
- [14] Timothy Ross “Fuzzy logic with engineering Applications”, McGraw Hill Edition 1997.