Comparison of Various Lacks of Cohesion Metrics

Taranjeet Kaur, Rupinder Kaur

Abstract-In software engineering there are plenty of applications used for reduced complexity and improved fault prediction approaches. In this paper we study various metrics that are not very much suitable to find fault classes in software. Basically using the concept of metrics to find fault classes and reduced complexity of classes. . various techniques like linear regression, logistic regression, one way ANOVA, principal component analysis, radial basis function network, support vector machines, single layer perceptron, multilayer perceptron, error correction learning, back propagation algorithm. all these techniques are used to find faulty classes and reduced complexity in software.

Keywords- Object oriented classes, class cohesion metrics, software quality, statistical approach.

I. INTRODUCTION

One of the fundamental concept of computer science is software metrics, In general language metric is a Mathematical term, it provide information for people to solve problems. Various metrics and tools are used to reduce complexity of software. Earlier we study various cohesion metrics like functional, logical, object, coincidental, layer, temporal, communication, sequence, procedure, some of these are weak metrics and some strong cohesion metrics. In this paper we study in which have lack of cohesion in metric LCOM,LCOM1,LCOM2,LCOM3,LCOM4,NLCOM, All these metrics are not find more faulty classes and also not very much suitable for reduced complexity. Main concern is that improved fault prediction power considers object oriented classes and also reduced complexity in software.
Main concern about metrics is

- Increasing fault prediction power.
- Reduced complexity.

II. RELATED WORKS

With the growth of technology, application of metrics in software engineering is also grow, Now a day's many area in computer science where metrics are used that software metrics are widely used in software engineering, first of all cohesion metrics is one of the important metric in software engineering and also it is a attribute of software quality, cohesion is concept like in which see that how methods of a class are closely related to one another in other words say that modules is highly cohesive, if module is highly cohesive means complexity of that classes reduced and that module is very reusable and maintainable. There are several different approaches to determine cohesion in object oriented system.

Manuscript received on February, 2013.
Taranjeet Kaur, research scholar have done B.Tech from Doaba institute of engineering and technology in kharar,now doing M.Tech CSE in lovely professional university,research area in software engineering, Phagwara, Punjab, India.
Rupinder Kaur, Assistant Professor ,have doneB.Tech from Baba Banda singh Bahadur engg college.Fatehgarh sahib(2010) M.Tech university institute of engg and tech. Punjab, India.

Various metrics and technique discussed here to find more faulty classes

LCOM1, LCOM2, LCOM3, LCOM4, LCOM5, NLCOM, ELCOM

Researcher chidamber and kemerer say about lack of cohesion metric is that metric in which number of methods not directly shares the attributes [11]

Another definition according to li and henry said that number of disjoint sets formed by intersection of m sets also say that LCOM is not very good to find fault classes that's why LCOM suggested not to be good cohesion metrics[11]
For example:

```
Class sampling class{
Int b1,b2;
Void m1()
{
b1=2;
}
Void m2()
{
b2=b1;
}
Void m3()
{
m2();
}
}
```

Figure (1)

Let m1,m2………mn be methods in a class.
Let I1,I2……In be number of instance variables in a class.
Let P set of non common instance variables.
P={(Ii,Ij)!Ii intersection Ij=phi}
Let Q set of all instance variables
Q={(Ii,Ij)!Ii intersection Ij is not equal to phi}
LCOM=number of noninstance variable minus number of instance variable (if number of noninstance variable is greater than number of instance variable); otherwise LCOM value is zero. According to figure 1 sampling class has three methods m1,m2 and m3 and integers are b1 and b2,method one has one variable b1 and value assign to b1 is 2,method 2 has variable b2,b2=b1,method 3 call method 2 .now calculate LCOM value on the basis of this figure. If LCOM value is high means good cohesion also we try to find low value thus we achieve high cohesiveness in the methods. In this figure LCOM value is two means two method pairs do not directly share the common attributes.
Chidamber and kemerer redefinition for LCOM1 is same as we discussed in above figure, LCOM 2 definition is disjoint set of methods in which two methods in same set are share instance variable [1]
For example:
Variables are: A, B, C, D
Functions are: M, N, O, P

Function M access variables {A,D}
Functions N access no variables.
Function O access variables {B,D}
Function P access variables {C}
According to this example calculate LCOM1 and LCOM2
Using LCOM1 formula:

Function M intersect N=null set
Function M intersects O= {D}
Function M intersect P=null set
Function N intersect O=null set
Function N intersect P=null set
Function O intersect P=null set
Number of methods that do not directly shares attributes whose value =5
Number of methods that directly share common attribute =1
LCOM1=4

Using LCOM2 formula:
According to definition of LCOM2 its
{M, O}, {N}, {P}
LCOM2=3

It would be derived from counting number of sets.
Another example of LCOM1 is that

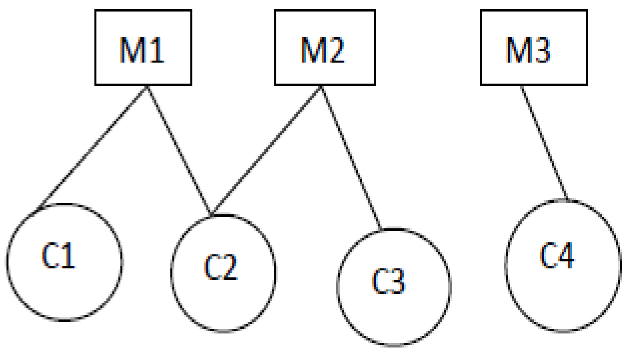


Figure (2)

In this case LCOM1 value is 2 and LCOM2 is 1.It is clear from this example lCOM2
is better than LCOM1 because LCOM2 has less value as compare to LCOM1.As compare to their range lCOM2 is also not good cohesion metrics its range lie in between[0,2].
After LCOM2, LCOM3 version is come

LCOM3
LCOM3 varies in range between [0, 1].LCOM3 indicates high cohesion and it is also a well defined class, it show simplicity of class and high reusability of class. A highly cohesive class provides high degree of encapsulation [2]
LCOM3 formula is
M: number of methods in a class
A: number of variables in a class
m.A: number of methods that access a variable.
Sum(m.A): number of methods over attributes
a: variables that shared may be or not

LCOM3=(m-Sum(mA)/a)/m-1
In other words also say that consider connected components of graph or also say that if we consider undirected graph G where vertices are methods of a class and edge between vertices if corresponding methods at least share one instance variable.

LCOM3 =|connected components of G|.

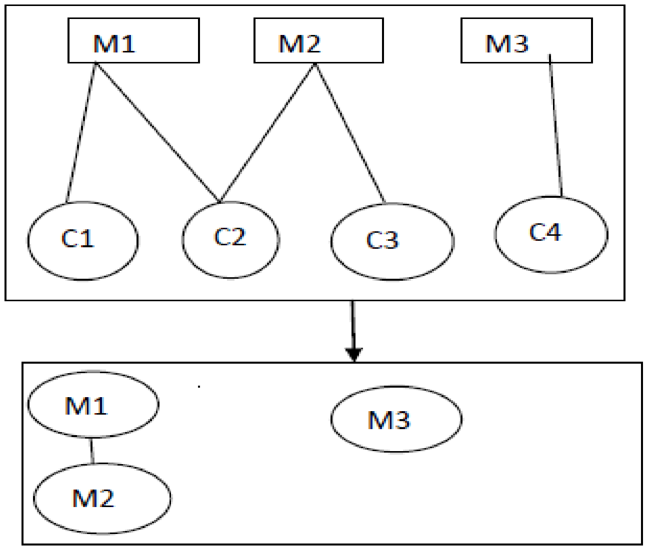
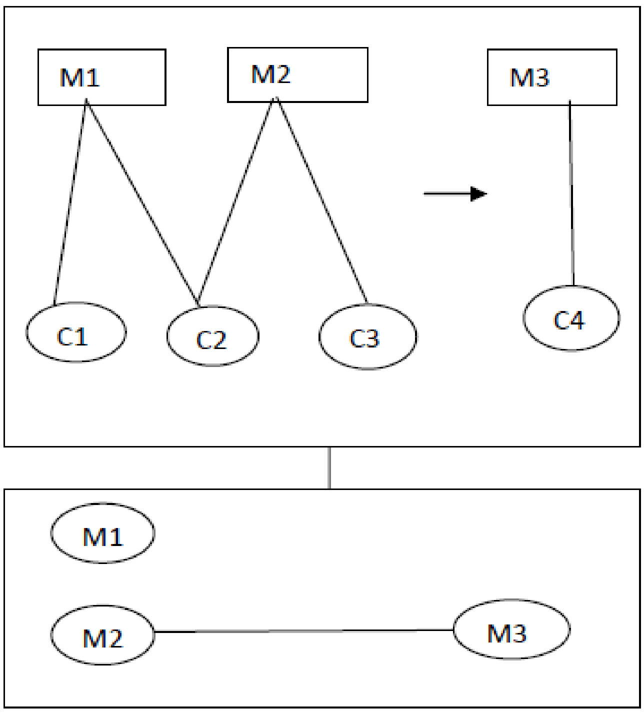


Figure (3)

LCOM4
It is like LCOM3 where graph G additionally has an edge between vertices representing methods Mi and Mj if Mi invokes Mj.
In other words also say that, it measure number of components in a class. a connected component is a set of related methods.
If LCOM4 =1 indicates a cohesive class, good class If LCOM4>=2 indicates a problem, class should be split into further smaller classes
If LCOM4=0 no methods in a class also called bad class.
Main concept related with LCOM4 is that, it is used for visual basic system as compare to earlier versions of LCOM metrics, like LCOM1, LCOM2, and LCOM3 these are not used for visual basic



Figure(4)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

LCOM5

LCOM5= 1/a∑μ(Aj)-m÷1-m
From the example
m=3, c=3
Thus LCOM5=2/3
For example:

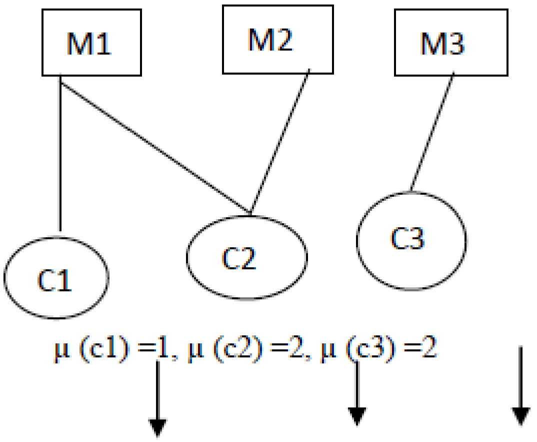


Figure (5)

.NLCOM

One problem with LCOM is that metric does not get normalized or standard values, thus NLCOM exist to overcome this problem, so different normalization techniques applied to get standard values.

Techniques are:
Sigmoid normalization
Bowless normalization
Best fit normalization

Best fit normalization to be best LCOM normalization approach, its range [0, 1] gives exact transformation for LCOM range [0, 1].

ELCOM

It is non information based and does not provide good result due to elimination of constructor.

TLCOM

it is non information based and compatible with only LCOM. According to definition it transitively or directly shares the attribute, thus this is better than LCOM. In the figure 1.TLCOM value is less as compare to LCOM value.

Table (1)

Metrics	Values
LCOM1	4
LCOM2	3
LCOM3	2
LCOM4	1
LCOM5	2/3
TLCOM	0

From table (1) show that TLCOM values less than LCOM1,LCOM2,LCOM3,LCOM4,LCOM5,means it is good cohesion as compare to other lCOM metrics, LCOM definition states that if value is high means bad cohesion if value is less means good cohesion.

III. PROPOSED WORK

Lack of cohesion is big issue in software metrics; *Jehad al Dallal* purposed TLCOM to improve LCOM. It takes simple classes to improve LCOM. Our purposed works is that take two open java software project and consider only inheritance classes and calculate values using regression model. Also check that LCOM improve if we consider Inheritance classes.

IV. CONCLUSION

This paper has attempted various cohesion based metrics from available literature. This is very helpful for developer and tester, mainly tester ,main problem is to find faulty classes, LCOM not very much good to find fault classes thus its various version come to find more faulty classes. Our future work is to find more faulty classes and reduced complexity in software using new techniques.

REFERENCES

[1]. Malkit singh and Dalwinder singh salaria (2012)"Approaches for software fault prediction "IJCST ,Vol 3.
[2]. Jehad al dallal(2010)"using TLCOM based OO lack of cohesion metrics",procedia computer science,Vol 3.
[3]. S.P.Sreeja And Dr.R.Sridaran(2012)"A survey on different approaches of determining cohesion based object oriented metrics"international journal of engineering research and development, Vol 4.
[4]. Ezekie okike(2010)"a proposal for normalized lack of cohesion in methods metrics using field experiment",IJCSI international journal of computer science issues,Vol(7)
[5]. Linda badri,Mourad badri and fadal tour(2011)"an empirical analysis of lack of cohesion metric for predicting testability of classes"international journal of software engineering and applications
[6] J.al.dallal(2008)"software similarity based functional cohesion metrics"journal of software engineering and applications
[7] Jehad.al dallas (2010)"mathematical validation of object oriented class cohesion metrics"international journal of computer science,IEEE,volume 4.
[8] J,bosnia, etzkorn c.davis and w.li"a class cohesion metrics for object oriented design"journal of object oriented programming.
[9]. S counsells t.hall d bowes(2009)"theoretical and empirical analysis of three slice based metrics for cohesion"
[10]. Shouki a ebad Moataz ahmed(2011)"evaluation framework for package level cohesion metrics"international conference on future information technology IACSIT volume 13
[11] letha etzkorn,carl davis and wei li"a practical look at lack of cohesion in methods metrics
[12] Ezikel okikie(2010)"a pedagogical evaluation and discussion about lack of cohesion in methods using field experiment.

AUTHOR PROFILE

Taranjeet kaur, research scholar have done B.Tech from Doaba institute of engineering and technology in kharar,now doing M.Tech CSE in lovely professional university,research area in software engineering, Phagwara ,Punjab, India.

Rupinder kaur, Assistant Professor ,have doneB.Tech from Baba Banda singh Bahadur engg college.Fatehgarh sahib(2010) M.Tech university institute of engg and tech. Punjab university ,Research area is software engg.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication