An Optimization Approach for Assesing and Ranking of Crowsourcing Software Components using Coupling and Cohision criteria

Nahid Karimi Mohammadabadi[1*], Mohammadreza Molahosseini Ardakani[2]

1. M.Sc of Computer Engineering, Faculty of Engineering, Meybod Branch, Islamic Azad University, Meybod, Iran.
2. Assistant Professor, Department of Computer Engineering, Faculty of Engineering, Ardakan University, Ardakan, Iran.
   Corresponding Author: Nahid Karimi Mohammadabadi

A B S T R A C T

Today, the use of Crowdsourcing as one of the techniques of participation in the project has been given special attention, and what has further increased the concept in various social activities was Web 2.0 technology. Collaboration is emerging as a new emerging business model that focuses on engaging the population in issues such as problem solving, participation in ideas, innovation, production, and service delivery processes, and has a direct impact on product quality, loyalty and customer satisfaction. One of the most important applications of multipart systems is the development of business models based on this method. In this research, the following were done through a collective: The idea of developing a Web-based software system 2 Developing some of the components of the idea idea. (In other words, after selecting the idea model from step 1 in colloquial, some of the components of this idea idea were given to six programmers. Then, the components generated by these six programmers were evaluated based on two criteria of connectivity and coherence and the components that have Lower connections and more coherence were selected. Given that the proposed method was a more appropriate criterion than the previous criteria, lcom and Gandhi's method, we proposed the proposed method as a selective criterion. In the following, the proposed method for the comparison of connectivity and coherence was compared with the methods of Lcom and Gandhi and its superior performance was shown in the evaluation. Based on the results, the proposed method was suitable for rating, in contrast to the Lcom and Gandhi methods that were suitable for coherence and connectivity evaluation. To better present the method, we first compared the class C created by the collaborators and 6 developers using the LCOM1 to LCOM4 methods to select the optimal class by these methods. Using the proposed method, we were able to rank and rank the connection between the collections that were the result of Crowdsourcing and accordingly the most connections to the DATABASE class that were called in almost all other classes and provide the connection of the classes with the model database Made.

INTRODUCTION

One of the latest efforts to control and use the power of the population in line with the company's goals Collective is called. Collaboration represents the action of a company or institution in outsourced work previously By employees is open to a large, undetermined network of people under an open call. This Action can be done in the form of generating a joint product, but often performed individually by individuals.

This chapter explores the generalities of research, including the basic concepts of Crowdsourcing, and the expression of the problem and the objectives of the research. We also examine the challenges in current methods and then express the hypotheses of the research and describe the structure of the research.

problem statement

According to Mr. Howe's definition, " Crowdsourcing action is the choice of an activity previously performed by an employee and referring it to a large and undefined group of people outside the organization during a public call" (Howe, 2006).

The term Crowdsourcing is a combination of two words \ "population \" and \ "outsourcing \". So colocation means outsourcing to the mass of people. Recruiting is a kind of outsourcing, but not to specific companies or organizations, but to large individuals through public calling, which is often the Internet. According to Soroviki (2114) and Nambinasa and Sohni (2112), the population

can be explained by a large group of unknown individuals. Awareness involves the notion that a company can not make its own population. Also, individuals are composed of a population with different backgrounds that includes people who are efficient and inexperienced. The Crowdsourcing is not limited to problem solving, not content creation (Shank, 2010).

In another definition, Crowdsourcing is a new emerging business model focusing on community participation in activities such as problem solving, the production and development of concepts such as participation in ideas, innovations, production, and service delivery processes that affect quality Product, loyalty and customer satisfaction have direct effect. (BIOS,.) 2111 The decentralization, informality, vertical and horizontality of communications, are not exclusive of information, including collective benefits. The business model \ " Crowdsourcing \" focuses on the voluntary participation of the population to produce innovative products (Frank, 2008).

The trend toward increasing the participation and application of customers by firms shows that firms have introduced new organizational options to increase their innovation and competitive advantage, and this reflects the importance Fundamental changes are in choosing enterprise business models, because it's important to know how the business can and should be mobilized by the use of Web technology during the commercialization process, in an effort to attract new ideas and finance many firms today. The expansion of knowledge management and outsourcing of creativity, and Crowdsourcing as one of the most important current trends in academic research in the field of entrepreneurship innovation and business models can be included in the business agenda, although during the commercialization process, firms can use existing types Crowdsourcing Improve the capability of thinking and financing but when using Crowdsourcing, the goal of participation and the conditions required to obtain desired results should be carefully considered (Habibi, Hamid Reza, 1391).

One of the most commonly used applications is the development of software. The use of a community in development Software is recognized by many companies in the world. Collective businesses will enable businesses to benefit from greater savings in operating costs and shorter time periods for more efficient, up-to-date and high-tech applications. IT managers and administrators use different methods to generate their software products. Design, development and testing can be done with a combination of methods for deploying forces within the organization, outsourcing to other software companies and mass production communities (Hassanzadeh, 1392).

On the other hand, the goal of software engineering is the development of high-quality software and tools for promoting high-quality software that is easy to use and maintain (Kalish et al., 2013). On the other hand, there are two important concepts in colocation. That Coupling and Cohesion are called.

Various researchers have been able to provide parameters for measuring the binding and cohesion components, including TCC, LCOM3, LCOM and DAC. All of these parameters have two common features: first, they express the relation between the two classes or the two methods with a binary value; secondly, they consider the relation of cohesion and connectivity as necessary (Joy and Scott, 2013).

A proposed method for assessing the connectivity and model coherence

Given the definitions of connectivity and cohesion, and using the models presented in this field, we can provide more appropriate metrics for evaluating these two components in object oriented programs. In order to provide an appropriate metric for evaluating the degree of connection between two classes, Definition of this metric. One of the most comprehensive definitions in this context is the definition of Joy & Scott:

Connectivity is an overlapping component of a system. Whether or not they interact with the components of a system. If they are very interdependent, applying changes to one will likely cause a lot of changes in the rest of the components (Joy and Scott, 2013).

What is to come from this definition is that not only calling classes by other classes increases the degree of connection of the program, but the common variables and attributes also increase this degree of connection because using a common variable in two classes causes the change in structure That variable, call classes also change. With a review of the metrics presented for the connection evaluation, it is a metric that coincides with Calls and common variables have not been addressed.

In addition, not only the call of a class increases the degree of connection, but also triggers the call But for evaluation of the whole system connection, only one of these calls is evaluated To be placed It means that it is not possible to obtain the degree of connection of the whole system from the sum of the degree of connection of all classes together.

The following (1) is suggested for calculating the connection grade of each class:

Degree of connection = number of calls by other classes + call count of other classes + number of common variables between two classes

Call Count by other classes = count in =CI

Number of calls to other classes= count out =Co

Number of common variables between two classes= max Count Attribute =MCA

DEGREE OF COUPLING = CO +CI+MCA

Given this relationship, the degree of connection (Table 1-3) for the model classes is as follows is:

Table 1. Connecting Model Classes in the Proposed Method

Class	Connection degree
A	26
B	6
D	4
E	7
F	6
G	5
H	5

According to Table 1-3, the highest degree of connection is in the Class A model. To calculate the coherence, we can also use the proposed method \ "A consolidated module module Of related activities that can not be split into sets. \ "

We can assess coherence based on two factors:

1-The number of variables used by a pair of methods. CA^1

2- he number of calls made by a method by another method in the same class CM^2

3- The number of pairs of methods that do not have a common variable minus the number of pairs of methods that have a common variable $p - q$

With the condition that:

$$\begin{cases} p\text{-}q, & \text{if } p\text{-}q >= 0 \\ 0, & \text{otherwise} \end{cases}$$

Therefore, for each class, these two factors need to be evaluated and aggregated together. The higher numbers indicate more coherence.

DEGREE OF COHESION = (P-Q)+AVERAGE(CA)-CM

It should be noted that this formula, like the Lcom formulas, measures incoherence, in other words, better than any The larger numbers will get less coherence.

Table 2 illustrates the coherence of proposed system classes:

Table 2. Coherence of proposed system classes

Class	Average number of variables used AVERAGE(CA)	The number of method calls by another method in cm class	P-Q	Degree of coherence
A	3	2	4	10
B	5	2	11	53
C	3	2	6	19
D	3	0	8	24
E	3	1	6	14
F	4	2	4	14
H	4	2	4	14

The highest degree of cohesion is in Class A and the lowest degree of integrity in Class B is observed. If the results of this table With the results of Table 2, it is clear that the proposed method better managed the coherence of the classes To specify. They chose both class B classes as the class with the least coherence, but Table 2 By providing the same number for several different classes, choosing the best coherence is difficult, if the method The pitch has made the difference better.

Selected classes were among the classes that different programmers (selectively by Crowdsourcing) Created. The selection criterion was for a high degree of cohesion and low connectivity. That had less priority with the connection Because the more

---

[1] Count attribute

[2] count method

interconnected classes reduce their portability and modularity, these two features have two important parameters Projects are done in a collective way.

For a better presentation of the methodology, let's take a sample of the selection of the developers' classes This section states:

When was the class of this class's product class, it was expressed here as the class c we do. The MCG graph before choosing this class is as follows:
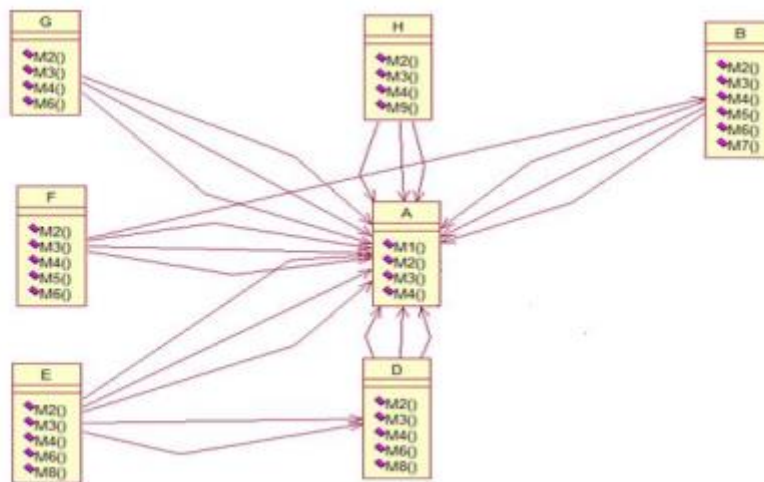
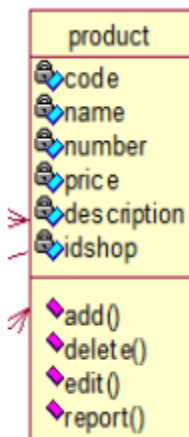The methods and variables of this class are given in figure 1, which should be coded by developers:

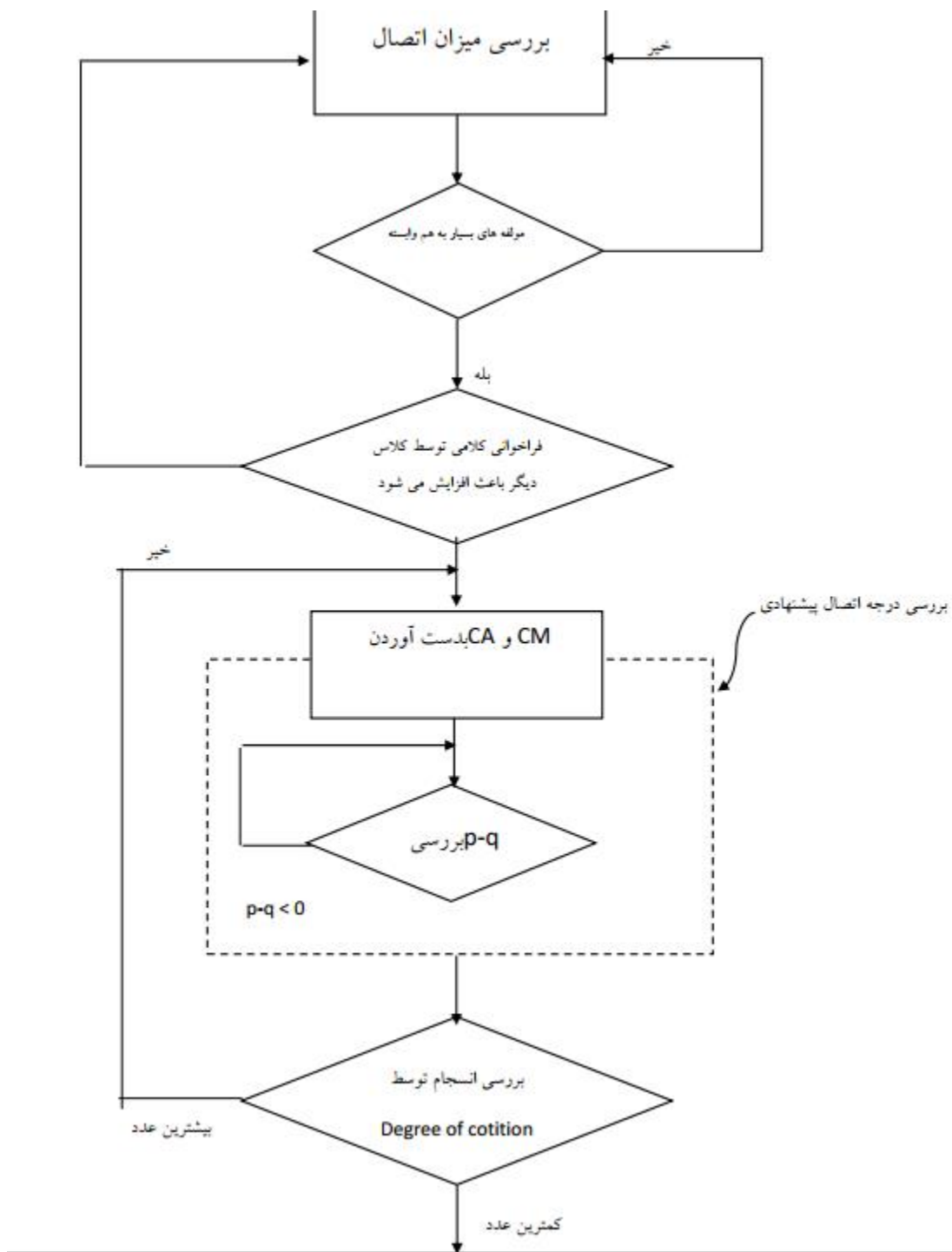Figure 2. Classes and class variables

بررسی میزان اتصال

خیر

مولفه های بسیار به هم وابسته

بله

فراخوانی کلامی توسط کلاس دیگر باعث افزایش می شود

خیر

بررسی درجه اتصال پیشنهادی

بدست آوردنCM و CA

بررسیp-q

p-q < 0

بررسی انسجام توسط
Degree of cotition

بیشترین عدد

کمترین عدد

Figure 3. Flowchart Proposed Method

Compare and Conclude

The use of collective recruitment in different areas increases the efficiency and globalization of the workforce One of the most important applications is the use of collaborative software development, analysis, design and testing. Software that is designed in this way will have more capabilities and performance, but this The software manufacturing process also has some problems, including the design of the design, There are many connections in the code generated by a developer, in other words better non-portability of the code Methods and classes produced by a developer. For this reason, it is necessary to have metrics

for Measurement of quality, reliability, portability, development and connectivity and the coherence of components derived from Crowdsourcing is considered. So far, in all of these areas, there are still some criteria There is room for improvement and more complete ideas.

Compare proposed method with other available methods

To provide a better method, first create a class C created by Crowdsourcing and 6 developers We compared the LCOM1 to LCOM4 methods to select the optimal class by these methods. Method LCOM1 metrics such as the number of shared variables and the number of pairs of methods that make a variable Do not subscribe. Of course, instead of these two, the number of methods in this method was critical. The LCOM2 method presents similar results with LCOM1, instead of the total number of methods A pair of methods that do not share a variable will sometimes cause numbers A more precise coherence is achieved, which is not so tangible. The LCOM3 method is the same as the previous two methods Only on the basis of the number of pairs of methods that share a variable, without paying attention to There are a number of these variables and methods calling each other (two factors in the proposed method in The evaluation was used. LCOM3 results are similar to those of the previous LCOM and only incoherence numbers It will be a little smaller. But in the LCOM4 method, the methods that call the other method are also important There is an item that optimizes the use of this method in new programming methods such as object-oriented.

The LCOM4 is better than the previous three methods, but for use in top class selection It does not have an application. This method only declares whether a class needs to be rewritten and Classes are divided into smaller classes (numbers greater than or equal to 2). There is no need to consider a method that is suitable for object-oriented programming, and both criteria And in addition to calling the methods each other, consider the number of invoked variables.

Also, among the methods presented for connecting a method that has complete evaluation parameters Can not find rankings Gandhi's method is one of the most suitable methods for evaluating connection and 112 Coherence, but this method, despite the fact that the ranking also does because of the low parameters of the case Evaluations usually provide equal numbers for methods and classes, and poses a difficult ranking task This method is appropriate for assessing connectivity and coherence, but does not have the required performance rating.

In this The research used a new idea to measure the connectivity and cohesion of classes, which is a hybrid proposal The famous methods were Lcom. A new idea is better than previous methods of connecting and coherence classes And can help in choosing the optimal components of colocation to help more than before Offer. The reason for this is because LCOM methods, especially LCOM1 and LCOM2, are all Coherent determination factors have not been taken into account. In other words, in LCOM methods only for the purpose of coherence assessment, the common use of variables in methods of a class was considered as determining factor of coherence. Considering the use of a method in another method as well as the number of variables Shared by methods that were considered in the proposed method, these two factors were also considered , Also in type lcom method The connection or, in other words, the number of variables that the connection between the two classes is shared with Not considered. For example, suppose a method in class A calls a method in class B. In the first case, this call is only shared by sharing a parameter and in second cases by sharing The parameter is done. Most of the connectivity measurement methods, the connection between Class A and B in both one-on-one modes They consider. But in the proposed method, the connection between the two classes in the second mode is more than the first one It turns out. Because to fix the connection and fix the dependence of the two classes and turn them into a modular class More changes are needed in the second case.

In the provided metric to evaluate the connection in addition to class calls that play an important role in raising Connection of components. Variables and common traits among classes were also interfered to contradict the Gandhi method And colleagues, only calls that do not determine the connection criterion. Moreover, unlike Gandhi's method, the number The class references are divided by the number of references to the class, and then the amount of connection is determined. We do this Summarized references and combined with common variables between classes.

In the Gandhi method, the number of connections is between 1 and 1, which is the maximum connection with a given number But in our method, the numerical connection is greater than zero and has no limits.

Table 3. Compare the advantages and disadvantages of the methods, lcom gandhi and the proposed method

Disadvantages	Advantages	Method name
It is only for measuring coherence. It does not do based on coherence Give	It performs integrity assessment well. It is appropriate for new programming (object oriented) methods.	Lcom4
Due to the low number of parameters being evaluated, they produce similar results and do not perform well.	Both for the assessment of coherence, it also binds evaluation and ranking Suitable for new programming (object oriented).	Gandhi

	It is also coherent to assess connectivity Also, the rating is done together Suitable for new programmers (object oriented) Given the low number of parameters evaluated Generate similar results and rank well It does not.	suggested method

Using the proposed method, we were able to get the connection between the classes that was the result of the task Based on this, most connections were made to the DATABASE class It was called in almost all other classes and provides the connection of classes with the model database Made. This component has methods for inserting information in the database, deleting them, and editing information That these operations were used in all other components. In addition, considering the attributes and The variables that each class stored in the database was different from the other classes, so a large number The shared variable is shared between this class and other classes, and this is the amount of connection this Classes with the other classes as much as possible.

According to the findings of the fourth chapter, after using the proposed method for assessing connectivity and coherence The codes written by six programmers for the class c containing 4 methods and 6 variables were the results contained in Table 2-2 was lost:

As shown in Table 1-2, lcom4 is the closest method to the proposed method for Coherence evaluation provides the same values for all methods, and therefore the work is the best choice The option of coherence is difficult. The reason for this is to pay attention to the number of common variables It does not evaluate merely by grouping methods that use the same variables.

That means, since all of our methods use six common variables in this method, they are grouped in one group So the number of groups is one. Of course, this is not a problem because this method is basically for Class ratings are not appropriate and are for classroom assessment only if they are consistent You need to make changes to their code so that they are coherent.

On the other hand, the Gandhi method, which is used to calculate the connection, is due to the low number of similar number assessment factors A lot of production is making it difficult to rank a ranking. This method, of course, is both for evaluation and for Ranking is due to the range of the response between 0 and 1 and the low score of the result evaluation factors It is not desirable.

REFERENCES

1. Bolat, T and Yılmaz, O. (2009), The relationship between outsourcing and organizational performance: Is it myth or reality for the hotel sector? International Journal of Contemporary Hospitality Management, Vol. 92 No. 2, pp. 7-92.

2. Brabham, D.C. (2009), Crowdsourcing as a Model for Problem Solving: An Introduction and Cases, Convergence: The International Journal of Research into New Media Technologies, 2 (2), pp.7-20.

3. Corney, J.R; Torres-Sánchez, C; Jagadeesan, A.P; Yan, X.T; Regli, W.C and Medellin, H,(2009),Putting the crowd to work in a knowledge-based factory, Advanced Engineering Informatics, N. 9, pp. 90-92.

4. Davis, J.G, (2009), From Crowdsourcing to Crowdservicing, Published by the IEEE Computer Society, pp. 24-29.