International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013

DOI: 10.7321/jscse.v3.n10.1

A Novel Technique of Optimization for Software Metric Using PSO

Anurag Kumar[1*], Amit Sinhal[2], Bhupendra Verma[3]

[1]Professor, Department of CSE, M. Tech Scholar
[2]Professor, Department of CSE, Technocrats Institute of Technology, Bhopal, India
[3]Professor Department of CSE
Email: [1]anurag_012@rediffmail.com, [2]amit_sinhal@gmail.com,[3] bk_verma3@gmail.com

Abstract. **Software Metrics present the key role of Software Development. Cost, Productivity and Quality are specific area of measurement in software field. The uncertainties which a bugs of effort estimation, researchers used the optimization to reduce it. Software cost estimation optimize, on basis of existing data set, in this paper we emphasize on COCOMO model with NASA18 data set. Software Effort Cost estimation is the process for measurement precisely the amount of effort required to complete the project. Regression rigorous method for estimation and Particle Swarm Optimization (PSO) is the austere method to work on the cost effort of software metricsin the modern era.**

Keywords: **NASA-18 Data Set, PSO, Regression, MMRE.**

* Corresponding Author:
Anurag Kumar,
Professor, Department of CSE, M. Tech Scholar, India
Email: anurag_012@rediffmail.com

I. Introduction

Software metrics refers to a broad range of measurements for computer software. Measurement can be applied to the software process with the intent of improving it on a continuous basis. Measurement can be used throughout a software project to assist in cost estimation, quality control, productivity assessment, and project control. Finally, measurement can be used by software engineers to help assess the quality of technical work products and to assist in tactical decision making as a project proceeds.
Software metrics provide a quantitative way to assess the quality of internal product attributes, thereby enabling the software engineer to assess quality before the product is built. Metrics for the analysis model focus on function, data, and behavior—the three components of the analysis model.
The Researcher to optimize the value of data, the use the COCOMO data set e.g. NASA 18 etc. [1]. . There are various techniques to optimize the value of the data set, regression, machine learning, analogy, genetic algorithm, data mining, neural network etc. [2, 3, 4]. PSO is another technique to optimize the value of a data set [9].
Particle swarm optimization algorithm is the simulation of birds feeding and social behavior, each particle has a memory and can remember to find the optimal solution in the process of finding the optimal solution obtained, and social behavior is that each particle to find the best solution of space, PSO compared with regression method. The Particle Swarm Optimization has no complex mating, mutation, natural selection, more simple and has faster convergence. In this paper we optimize the result with the help of PSO and their different model [5] [6]. We also use the logarithmic equation of regression and Compare the result.

II. Literature Review

There are various work on effort estimation some work are discussed here .Function Point Based to estimate the effort to influence on the size and factor [7]. A new approach based on fuzzy logic, linguistic quantifiers and analogy based reasoning is proposed to enhance the performance of the effort estimation in software projects dealing with numerical and categorical data [8].The basic input for the effort estimation is size of project. A number of models have been proposed to construct a relation between software size and Effort; however we still have problems in effort estimation because of uncertainty

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

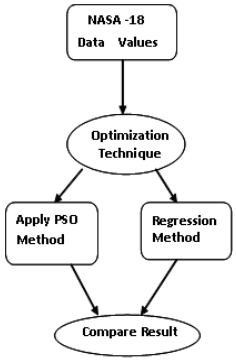
existing in the input information [9]. More work has been done on software effort estimation. Genetic approach is also an optimization technique based on different model in the paper we studied that GA base result compare with another effort model [10]. Fuzzy method also used for effort estimate. The Gaussian method applies through fuzzy is effective role in estimation [11]. The Analogy-X approach is a set of procedures that utilize the principles of the Mantel randomization test to provide inferential statistics to Analogy, it provides a further empirical evaluation of Analogy-X uses different kinds of datasets [12]. Case-Based Reasoning (CBR) approach integrated with multi-agent technology to retrieve similar projects from multi-organizational distributed data sets. The study explores the possibility of building a software cost estimation model by collecting software cost data from distributed predefined project cost database [13].
In the [14] paper introduce, the aspect of feature subset selection by using a generic backward input selection wrapper is investigated .In this literature, a model which combines genetic algorithm (GA) with support vector machines (SVM). We can find the best parameter of SVM regression by the proposed model, and make more accurate prediction. In this paper test and verify the model by using the historical data in COCOMO [15].
A neural network is a massive parallel distributed processor made up of simple processing units, which has a natural property for storing experimental knowledge and making it available for use. It associates the brain in two respects. Knowledge is acquired by the network from its environment through a learning process and Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. In this literature, researcher to present the two network models used for the case study i.e. Radial Basis Neural Network and Generalized Regression Neural Network [16].

III .Regression and PSO

Linear regression uses the fact that there is a statistically significant correlation between two variables to allow you to make predictions about one variable based on your knowledge of the other. For linear regression to work there needs to be a linear relationship between the variables. The regression line is a kind of "moving average" that is drawn through the balancing point between the dots at each point on your X-axis. [17]. Regression can be used for prediction, estimation, and hypothesis testing, and modeling causal relationships.
Particle swarm optimization (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA) [18]. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles [19].

IV. Proposed Work and Methodology

In this paper we follow that PSO technique is a good technique to resolve the uncertainty of data set and optimize the value corresponding to the effort. Regression also good technique but it takes time and expensive .PSO is relevant to show the result and with less time .In the paper we compare the result using PSO. Accuracy of parameters is much better than Regression generated. Because it will take a maxim mum time and in its no need the predict values in a procedural way in PSO[20] , minimum fitness is the initial value to start to optimize value these value having to type pbest and gbest .And a collection of the swarm iterate till the best show required fitness value. Each particle tries to modify its current position and velocity according to the distance between its current position and pbest, and the distance between its current position and gbest. In figure 1.1 Data Set optimize the value with PSO and Regression and find the result.



METHODOLOGY (ALOGORITHM)

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

Input: Size of Software Projects, Measured Efforts, Methodology. Output: Optimized Parameters for Estimating Effort. The following is the methodology used to tune the parameters in the proposed model for Software Effort Estimation. [21] [22]
Step 1: Initialize "n" particles with random positions Pi and velocity vectors Vi of tuning parameters. We also need the range of velocity between [-Vmax, Vmax]. The Initial positions of each particle are personally Best for each Particle [23].

Step 2: Initialize e the weight function value w with 1 and weighting parameters cognitive learning factor c1, social coefficient c2 with 2.0

Step 3: Repeat the following steps 4 to 9 until the number of iterations specified by the user or Particles Exhaust.

Step 4: for i = 1,2,……, n do //
For all the Particles For each particle position with values of tuning parameters, evaluate the fitness function. The fitness function here is Mean Absolute Relative Error (MARE). The objective in this method is to minimize the MARE by selecting appropriate values from the ranges specified in step 1.

Step 5: Here the Pbest is determined for each particle by evaluating and comparing measured effort and estimated effort values of the current and previous parameter values.
If fitness (p) better than fitness
(Pbest) then: Pbest = p.

Step 6: Set the best of 'Pbest' as global best – Gbest. The particle value for which the variation between the estimated and measured effort is the least is chosen as the Gbest particle.

Step 7: Update the velocity and positions of the tuning parameters with the following equations
For j = 1, 2,…… m do // For number of
Parameters, our case m is 2or 3 or 4
Begin

$$v_{ij}^{t+1} = v_{ij}^{t} + c_1 r_{1j}^{t}\left[P_{best,i}^{t} - x_{ij}^{t}\right] + c_2 r_{2j}^{t}\left[G_{best} - x_{ij}^{t}\right] \quad \ldots\ldots\ldots (5)$$

$$S_i^k = V_i^{k+1} + S_i^k \quad \ldots\ldots\ldots\ldots (6)$$

End;
Step 8: Give the Gbest values as the optimal
Solution.
Step 9: Stop

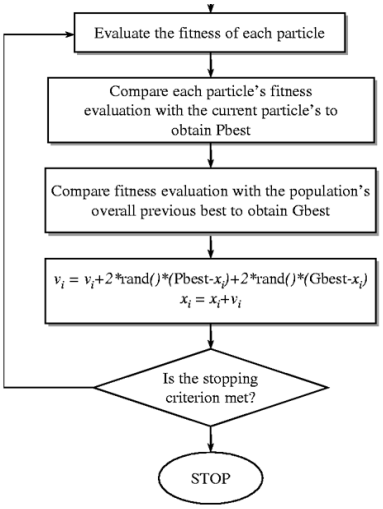


Figure 1.2

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

Above figure 1.2 shows well defined step of PSO process.It follows as evaluate the fitness of each particle ,fitness obtain pbest particle,now comparefitness with population and obtain gbest . These process continue to evaluate the gbest position.

VI .Performance Indicators

It's usually used Mean of MRE (MMRE) and Prediction level (Pred) as an accurate reference value in the research of software effort estimation. In this study, using the Pred and MMRE as accuracy reference value.

MMRE Software effort estimation in the assessment of evaluation criteria commonly used Mean Magnitude of Relative Error (MMRE), the formula as equation below. [24][25]

$$\text{MRE}=\sum_{i=1}^{15}\frac{actual\ effort-measured}{actual\ effort}*100$$

$$\text{MMRE}=\frac{100}{N}\sum_{i=1}^{15}\frac{actual\ effort-measured}{actual\ effort}$$

In this study, MMRE for the PSO algorithm as the effort estimates of fitness value and evaluation criteria. MMRE value is the smaller that the prediction effort closer the actual effort. Which actual_effort is actual effort, predicated _effort is prediction effort N is the number of projects i is NO. i project.

V .Result and Analysis

In this paper we take NASA18 data for optimization .In this data having size, effort and one factor me. The table given below: table 1.1

Size	Effort	ME
90.2	115.8	30
46.2	96	20
46.5	79	19
54.5	90.8	20
31.1	39.6	35
67.5	98.4	29
12.8	18.9	26
10.5	10.3	34
21.5	28.5	31
3.1	7	26
78.6	98.7	35
9.7	15.6	27
12.5	23.9	27
101	138.3	34
4.2	9	19
7.8	7.3	31
2.1	5	28
5	8.4	29

Table 1.1

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

For Implementation, Java is suitable for various inbuilt method random value function is also a number generator .We calculate the linear regression method for obtaining the result. This regression method is calculating the predicted value of the equation:
Log (Ef) = logA+ (logSize) *B …… (4)
Below table 1.2 shows the result of regression and tell the value of parameter A=1.127462, B=1.028.

S.No	Size	Actual Effort	Regression
1	90.2	115.8	0.28803
2	46.2	96	0.28128
3	46.5	79	0.274302
4	54.5	90.8	0.279283
5	31.1	39.6	0.249881
6	67.5	98.4	0.282167
7	12.8	18.9	0.224243
8	10.5	10.3	0.203597
9	21.5	28.5	0.238414
10	3.1	7	0.190642
11	78.6	98.7	0.282276
12	9.7	15.6	0.217678
13	12.5	23.9	0.232321
14	100.8	138.3	0.294454
15	4.2	9	0.199055
16	7.8	7.3	0.192043
17	2.1	5	0.179469
18	5	8.4	0.19674
		MMRE	0.239215278

Table 1.2

Now our experiment through Particle swarm intelligence (PSO) Model-1 & Model -2. In Model -1 we use 10 swarms of values and optimize the swarm again, they obtain values are in following manner. We use the velocity v [-5,5] range of a [0,10] and b [-1,1].
Now in our model the parameters are tuned using the above PSO methodology .The Update of velocity and positions of Parameter "a" is

$$v_{aj}^{t+1} = v_{aj}^{t} + c_1 r_{1j}^{t}\left[P_{best,i}^{t} - x_{ai}^{t}\right] + c_2 r_{2j}^{t}\left[G_{best} - x_{ai}^{t}\right]$$

$$x_{ai}^{t+1} = x_{ai}^{t} + v_{ai}^{t+1}$$

Updating the velocity and get the value of b

$$v_{bj}^{t+1} = v_{bj}^{t} + c_1 r_{1j}^{t}\left[P_{best,i}^{t} - x_{bi}^{t}\right] + c_2 r_{2j}^{t}\left[G_{best} - x_{bi}^{t}\right]$$

$$x_{bi}^{t+1} = x_{bi}^{t} + v_{bi}^{t+1}$$

$$v_{cj}^{t+1} = v_{cj}^{t} + c_1 r_{1j}^{t}\left[P_{best,i}^{t} - x_{ci}^{t}\right] + c_2 r_{2j}^{t}\left[G_{best} - x_{ci}^{t}\right]$$

$$x_{ci}^{t+1} = x_{ci}^{t} + v_{ci}^{t+1}$$

Table 1.3 shows the Model –I A=1. 4131, B=0. 9845 Parameter value of Model-II
A=0.6178, B=1.1369, C=0.2601

S.No	Size	Actual Effort	ME	Model-I	Model-II

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

1	90.2	115.8	30	0.026515	0.06741
2	46.2	96	20	0.359172	0.464238
3	46.5	79	19	0.216294	0.346473
4	54.5	90.8	20	0.202796	0.323817
5	31.1	39.6	35	0.052205	0.081557
6	67.5	98.4	29	0.091913	0.19854
7	12.8	18.9	26	0.08006	0.186062
8	10.5	10.3	34	0.388982	0.398916
9	21.5	28.5	31	0.016515	0.11618
10	3.1	7	26	0.385077	0.084033
11	78.6	98.7	35	0.051718	0.049139
12	9.7	15.6	27	0.151748	0.197852
13	12.5	23.9	27	0.289306	0.362095
14	100.8	138.3	34	0.041135	0.114039
15	4.2	9	19	0.35506	0.310107
16	7.8	7.3	31	0.462572	0.556324
17	2.1	5	28	0.413284	0.186673
18	5	8.4	29	0.179593	0.012798
			MMRE	0.209108	0.225347

Table 1.3

In this section the figure 1.3 shows the graphical representation of regression, model-I and model-II of effort of each nasa18 datum. Figure 1.4 shows the MMRE of each model, figure1. 6 shows that the MMRE value reduces by PSO is fine.

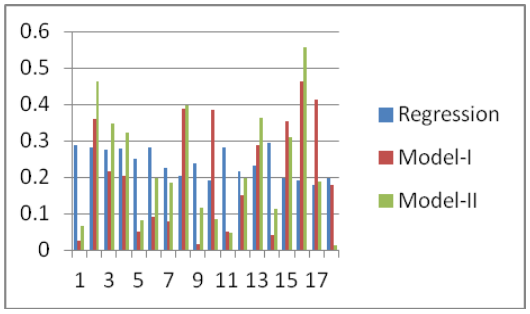


Figure 1.3

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

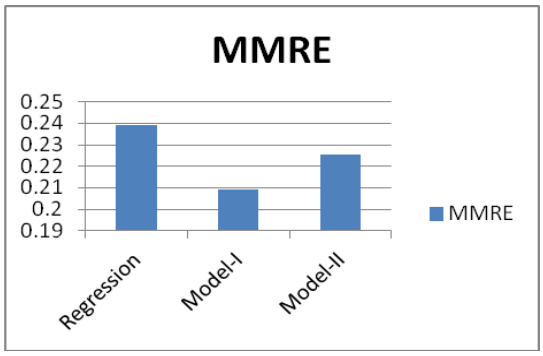


Figure 1.4

VII CONCLUSION

In this paper, we analysis the result between regression and PSO. We found that PSO provides an efficient way to optimize the effort prediction. The linear regression method also gives fine results but it will to time consuming. Regression is well and good for small problem. Uncertainty of data is not easily reduced by regression so Particle Swarm Optimization sufficient to reduce the uncertainty in the data set. In future we can use another techniques of Swarm Intelligence in this methodology.

REFERENCE

[1] B. Boehm, E. Horowitz, R. Madachy, D. Reifer, B.K. Clark, B. Steece, A.W. Brown, S.Chulani, and C. Abts, "Software Cost Estimation with Cocomo II", Prentice Hall, 2000.
[2] Tim M enzies, Zhihao Chen, Jairus Hihn, and Karen Lum, "Selecting Best Practices for Effort Estimation", IEEE Transactions On Software Engineering, November 2006.
[3] Alaa F. Sheta, "Estimation of the COCOMO Model Parameters Using Genetic algorithms for NASA Software Projects", Science Publications , 2006.
[4] Jin-Cherng Lin Han-Yuan Tzeng, "Applying Particle Swarm Optimization to Estimate Software Effort by Multiple Factors Software Project Clustering" IEEE 2010 .
[5] Kennedy, J. And Eberhart, R. C. , "Particle swarm optimization". Proc. IEEE Int'l. Conf. on Neural Networks, IEEE Service Center , 1995 .
[6]. Matthew S, An Introduction to Particle Swarm Optimization, Department of Computer Science, University of Idaho, pp.1-8 November 7, 2005
[7] Dowming Yeh , Deron Wang,Shu-Lan Chu, "A Specific Effort Estimation Method Using Function Point", Journal Of Information Science And Engineering 27, 1363-1376 (2011)
[8] S. Malathi , Dr. S. Sridhar, "Estimation Of Effort In Software Cost analysis For Heterogenous Dataset Using Fuzzy Analogy" (IJCSIS) International Journal of Computer Science and Information Security, 2012.
[9] Prasad reddy.P.V.G.D, Ch.V.M.K.Hari, "Software Effort Estimation Using Particle Swarm Optimization with Inertia Weight", Global Journal of Computer Science and Technology, October 2011.
[10] Kavita C , GA based Optimization of Software Development effort estimation, IJCST, Sep -2010.
[11]S.Malathi1 , Dr.S.Sridhar " A Classical Fuzzy Approach for Software Effort Estimation on Machine Learning Technique" IJCSI International Journal of Computer Science Issues, November 2011.
[12] J.Keung,"Empirical evaluation of analogy-x for software cost estimation", in ESEM '08: Proceedings of the second ACM-IEEE international symposium on Empirical engineering and measurement. New York,2008.
[13]Hasan Al-Sakran "Software Cost Estimation Model Based on Integration of Multi-agent and Case-Based Reasoning" Journal of Computer Science Publications in, 2006.
[14] Prasad Reddy P.V.G.D, Sudha K.R, Rama Sree P and Ramesh S.N.S.V.S.C "Software Effort Estimation using Radial Basis and Generalized Regression Neural Networks" Journal Of Computing, , MAY 2010.
[15]Karel Dejaeger, Wouter Verbeke,David Martens,Bart Baesens "Data Mining Techniques for Software Effort Estimation: A Comparative Study"IEEE Software Engineering, March-April 2012 .
[16] Jin-Cherng Lin , Chu-Ting Chang, Sheng-Yu Huang "Research on Software Effort Estimation Combined with Genetic Algorithm and Support Vector Regression", Computer Science and Society (ISCCS), 16-17 July 2011.

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.3,No.10, 2013
Published online: Nov 25, 2013
DOI: 10.7321/jscse.v3.n10.1

[17] Marian Petre, David Budgen and Jean Scholtz, "Regression Models of Software Development Effort Estimation Accuracy and Bias", Empirical Software Engineering, 9, 297–314, Kluwer Academic Publishers. Manufactured in The Netherlands, 2004.
[18] CH. V. M. K. Hari, Tegjyot Singh Sethi, B. S. S. Kaushal, Abhishek Sharma, "CPN-A Hybrid Model for Software Cost Estimation",,IEEE 2011.
[19] Srinivasa Rao T. , Hari CH.V.M.K. (2) and Prasad Reddy P.V.G.D , "Predictive and Stochastic Approach for Software Effort Estimation", Int. J. of Software Engineering, IJSE Vol. 6 No. 1 January 2013.
[20]. Felix T. S. Chan and Manoj Kumar Tiwari, (2007), Swarm Intelligence: Focus on Ant and Particle Swarm Optimization, I-TECH Education and Publishing, 2007.
[21] S.A. Andreas, P. Efi, and S. Christos, Evolving Conditional Value Sets of Cost Factors for Estimating Software Development Effort, 19th IEEE International Conference on Tools with Artificial Intelligence, 2007.
[22] Bingchiang Jeng, Dowming Yeh, Deron Wang, Shu-Lan Chu "A Specific Effort Estimation Method Using Function Point", Journal Of Information Science And engineering ,2011
[23] Sultan Aljahdali, Alaa F. Sheta, "Software Effort Estimation by Tuning COOCMO ModelParameters Using Differential Evolution", Computer Systems and Applications (AICCSA), IEEE/ACS International Conference on 16-19 May 2010.
[24] Prasad Reddy ,Hari CH.V.M.K.,Srinivasa Rao T., "Multi Objective Particle Swarm Optimization for Software Cost Estimation", International Journal of Computer Applications, October 2011
[25]. Wu B, Zheng Y, Liu S, and Shi Z, CSIM: A Document Clustering Algorithm Based on Swarm Intelligence, DOI:0-7803-7282-4/02@IEEE, 2002.