Parameter Optimization for Software Metric using Particle Swarm Optimization

Anurag Kumar
Senior Lecturer
Dept of CSE
IET, Bundelkhand University, India

Akshay Saxena
Senior Lecturer
Dept of CSE
IET, Bundelkhand University, India

Hemant Rai
Senior Lecturer
Dept of CSE
IET, Bundelkhand University, India

ABSTRACT
Software Metrics have an important role in Software Development. Cost, Productivity and Quality are specific area of measurement in software metrics. Parameter optimization is great challenge in software metrics. Scientists have used various techniques to optimize the parameter like as Artificial Intelligence, Neural Network and Genetic Algorithm etc. In this thesis, Particle Swarm Optimization (PSO) is proposed as optimization technique. PSO algorithm is a multi-agent parallel search technique which maintains a swarm of particles and each particle represents a potential solution in the swarm. Therefore this austere method is used to work on the parameter optimization in software metrics. An approach of two model structure of PSO has been used for optimizing the parameter. Standard NASA-18 data set is used to evaluate the proposed approach. PSO based models show better result as compared to regression method.

Keywords
Particle Swarm optimization(PSO), Regression, Mean Magnitude Relative Error(MMRE), NASA 18 data set.

1. 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, these three components are analyzing the model.

The Researcher to optimize the value of data, the use the COCOMO data set that is NASA 18 [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 obtaining he result with the help of PSO and their different model [5, 6].

2. 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 [9]. More work has been done on software effort estimation. Genetic approach is also an optimization technique based on different model 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). The parameter find using SVM regression model, and make more accurate prediction. COCOMO Data test and verify the model by using the historical data in [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 that is , 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, researchers to present the two network models used for the case study i.e. Radial Basis Neural Network and Generalized Regression Neural Network [16].

3. 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. [17]. Regression

also a technique for optimizing the parameter. 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].

4. PROPOSED WORK METHODOLOGY

In this paper, the parameter optimization is done by PSO. PSO technique optimizes the value corresponding to effort. PSO has a limited number of parameters and the impact of parameters to the solutions is small compared to other optimization techniques. In PSO, particles begin searching between each decision variable's initial boundary values, these are not necessarily boundary constraints since particles are generally allowed to search outside of each decision variable's range of values. PSO use the initial boundary values as boundary constraints to prevent particles from exploring outside a fixed search space. In PSO [20], minimum fitness is the initial value to start to optimize value these value having to type pbest and gbest and collection of the swarm iterate till the best value 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.

Here two models are describing to work with PSO. These models are used Boehm's basic model and it's equation. In figure 1 show that input data optimize the value with PSO and Regression and compare their result.

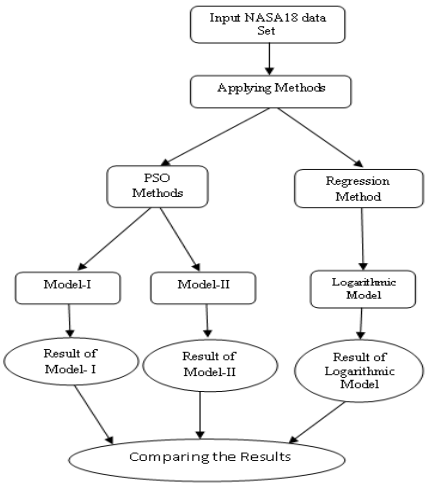

Figure 1: Proposed Work Model

Pso Methodology (Alogorithm)

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. For generating the random value takes 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

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

$$x_{ij}^{t+1} = v_{ij}^{t+1} + x_{ij}^{t} \quad \ldots\ldots\ldots (2)$$

End;

Step 8: Give the Gbest values as the optimal solution.

Step 9: Stop

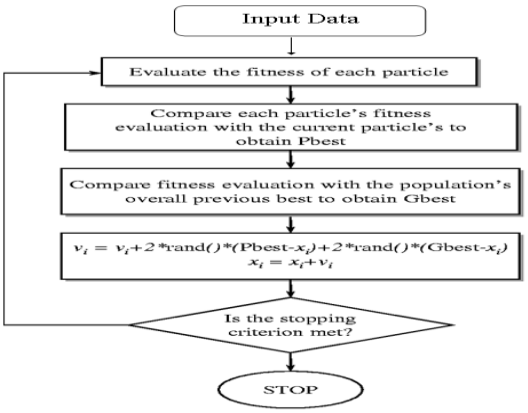

Figure 2: Data Flow chart of PSO

Above figure 2 shows well defined step of PSO process. It follows as evaluate the fitness of each particle, fitness obtain pbest particle, now compare fitness with population and obtain gbest. These processes continue to evaluate the gbest position.

5. 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 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 (3) (4) below [24, 25].

$$\text{MRE}=\sum_{i=1}^{N}\frac{\text{actual effort} - \text{measured}}{\text{actual effort}} \quad \ldots\ldots\ldots (3)$$

$$\text{MMRE}=\frac{1}{N}\sum_{i=1}^{N}\frac{\text{actual effort} - \text{measured}}{\text{actual effort}} \quad \ldots\ldots (4)$$

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..

6. CONCLUSION

In this paper, an approach of two model structures to optimize the parameter for the software effort using PSO. The performances of the developed models were tested on NASA 18 software project. The developed models were able to provide good estimation capabilities. From the implementation of the results, it is observed that the used method has effectively estimated the average effort for the software project datasets in comparison to regression method. It analyze that the use of PSO technique to build suitable model structure for the software effort. PSO can also use significantly in different domain.

7. REFERENCES

[1] B. Boehm, E. Horowitz, R. Madachy, "Software Cost Estimation with Cocomo II", Prentice Hall, 2000

[2] Tim M enzies, Zhihao Chen, Jairus Hihn, Karen Lum, "Selecting Best Practices for Effort Estimation", IEEE Transactions On Software Engineering, Vol. 32, No-11, pp 883-895,November 2006..

[3] Alaa F. Sheta, "Estimation of the COCOMO Model Parameters Using Genetic algorithms for NASA Software Projects", Science Publications, Vol-2, No-2 pp 118-123, in 2006.

[4] Jin-Cherng Lin Han-Yuan Tzeng, "Applying Particle Swarm Optimization to Estimate Software Effort by Multiple Factors Software Project Clustering" IEEE , pp 1039 – 1044, 2010 .

[5] Kennedy, J. and Eberhart, R. C. , "Particle swarm optimization", Proceeding IEEE International Conference on Neural Networks, IV,1942–1948. Piscataway, 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 in , Vol-27, No. 4, pp. 1363-1376, 2011.

[8] S. Malathi, Dr. S. Sridhar, "Estimation Of Effort In Software Cost analysis For Heterogeneous Dataset Using Fuzzy Analogy" International Journal of Computer Science and Information Security, Vol.10, No-10, 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, Vol. 11, Issue-18, October 2011.

[10] Kavita C , "GA based Optimization of Software Development effort estimation, International Journal of Computer Science and Technology", Vol.1, Issue-1, pp 38-40, September -2010.

[11] S.Malathi1, Dr.S.Sridhar, "A Classical Fuzzy Approach for Software Effort Estimation on Machine Learning Technique" , International Journal of Computer Science , Vol. 8, Issue-6, pp 249-253, November 2011.

[12] J.Keung, "Empirical evaluation of analogy-x for software cost estimation", Proceedings of the second ACM-IEEE international symposium on Empirical Engineering and Measurement, pp 294-296, 2008.

[13] Hasan Al-Sakran "Software Cost Estimation Model Based on Integration of Multi-agent and Case- Based Reasoning" Journal of Computer Science, Vol-2, pp 276-282, 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, Vol-2, Issue -5, pp 85-92, May 2010.

[15] Karel Dejaeger, Wouter Verbeke,David Martens,Bart Baesens "Data Mining Techniques for Software Effort Estimation: A Comparative Study"IEEE Software Engineering, in 2012 .

[16] Jin-Cherng Lin ,Chu-Ting Chang ; Sheng-Yu Huang "Research on Software Effort Estimation Combined with Genetic Algorithm and Support Vector Regression", International Symposium on Computer Science and Society, pp 349 – 352, July 2011.

[17] Marian Petre, David Budgen and Jean Scholtz, "Regression Models of Software Development Effort Estimation Accuracy and Bias", Empirical Software Engineering, Kluwer Academic Publishers, Vol. 9, pp 297-314,2004.

[18] CH. V. M. K. Hari, Tegjyot Singh Sethi, B. S. S. Kaushal, Abhishek Sharma, "CPN-A Hybrid Model for Software Cost Estimation", IEEE, pp 902 – 906, 2011.

[19] Srinivasa Rao T., Hari CH.V.M.K. and Prasad Reddy P.V.G.D, "Predictive and Stochastic Approach for Software Effort Estimation", International Journal of Software Engineering, Vol.-6, No-1, pp 97-115,January 2013.

[20] Felix T. S. Chan and Manoj Kumar Tiwari, "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, Vol-17, No-4, pp1363-1376,2011.

[23] Sultan Aljahdali, Alaa F. Sheta, "Software Effort Estimation by Tuning COOCMO Model Parameters Using Differential Evolution", Computer Systems and Applications, IEEE/ACS International Conference, pp 1-6, 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 Volume 32– No.3, pp 13-17, October 2011.

[25] Wu B, Zheng Y, Liu S, and Shi Z, CSIM: A Document Clustering Algorithm Based on Swarm Intelligence, IEEE, 2002.