

Enhancement of Multi-Objective Particles Swarm Optimization Using Clustering Techniques

¹Amita Talyan, ²Anoop kumar, ³Rajan sachdeva

¹Master in CSE, GGS Kharar (PTU)

²Master in IT, LPU

³Master in CSE, GGS Kharar (PTU)

Abstract: Planning is the most vital part of project management. It describes the resources that we need to complete the project successfully. Software cost estimation is a part of planning. It describes the estimated cost and time required to complete the project. The input in software cost estimation is the size of the code and cost drivers. The output is the Effort in terms person per month. Our proposed model is for tuning parameters of COCOMO model software cost estimation using Multi Objective (MO) Particle Swarm Optimization (PSO). We will be using clustering methods to divide the data items into number of clusters and PSO be used then for parameter tuning of each cluster. The clusters and the tuned parameters will be trained on Neural Network by back propagation algorithm. The results will be compared for the improvement of the previous work.

Keywords: Mutiojective (MO) Particle Swarm Optimization (PSO), integrated development environment (IDE), Constructive cost model (COCOMO), SWARM intelligence.

1. INTRODUCTION

Software cost estimation can be defined as the estimated judgment of the expenses for an undertaking. Fetched estimation will never be a definite science on the grounds that there are an excess of variables included in the figuring for an expense assessment, for example, human, specialized, ecological, and political. Furthermore, any process that includes a critical human variable can never be correct in light of the fact that people are excessively unpredictable to be completely unsurprising. Moreover, programming advancement for any reasonable estimated task will unavoidably incorporate various assignments that have complexities that are hard to judge on account of the many-sided quality of programming frameworks.

Cost estimation is typically measured regarding effort. The most widely recognized metric utilized is individual months or years (or man months or years). The effort is the measure of time for one individual to work for a certain time of time. It is essential that the particular attributes of the advancement environment are requiring into record when contrasting the effort of two or more undertakings in light of the fact that no two improvement situations are the same. An acceptable sample of contrasts being developed situations are the measure of time individuals work in distinctive nations; the run of the mill workweek in North America is 40 hours every week, while in Europe the regular workweek is 35 hours every week (Londeix, 1987). Subsequently, when contrasting a venture from North America and a task from Europe, a transformation variable would need to be utilized for a precise correlation. Diverse variables can be utilized for expense estimation, which prompts a trouble when looking at undertakings if standard models or apparatuses are not utilized. Case in point, an expense appraisal can incorporate elements from administration, advancement (e.g., preparing, quality certification), and different zones particular to an association.

Cost estimation is an essential apparatus that can influence the arranging and planning of a task. Since there are a limited number of assets for an extend, the majority of the gimmicks of a prerequisites archive can regularly not all be incorporated in the last item. An expense gauge done toward the start of an unde/rtaking will help figure out which peculiarities can be incorporated inside the asset stipulations of the task (e.g., time). Necessities can be organized to guarantee that the most critical gimmicks are incorporated in the item. The danger of an undertaking is decreased when the most vital gimmicks are incorporated toward the starting on the grounds that the intricacy of a task increments with its size, which implies there is more open door for oversights as advancement advances. Hence, Cost estimation can have a huge effect on the life cycle and calendar for a task.

Cost estimation can likewise have a vital impact on asset allotment. It is judicious for an organization to distribute better assets, for example, more experienced staff, to immoderate activities. Labor stacking is a term used to gauge the quantity of designing and administration faculty allotted to an undertaking in a given measure of time. The majority of time, it is more awful for an organization if an exorbitant venture falls flat than if a less immoderate task comes up short. At the point when devices are utilized for estimation, administration and engineers can even explore different avenues regarding exchanging off a few assets (or components) with others while keeping the expense of the undertaking consistent. Case in point, one tradeoff may be to put resources into an all the more compelling integrated development environment (IDE) with the goal that the quantity of faculty chipping away at an undertaking could be lessened. Cost estimation has an expansive effect on undertaking arranging and administrat.

2. LITERATURE REVIEW

G. Sivanages Rao, CH.V. Pani Krishna and K. Rajasekhara Rao: et al worked on COCOMO and MARE and used Multi Objective PSO to tune the parameters. The parameters of model tuned by using MOPSO considered two objectives, Mean Absolute Relative Error and Prediction. They observed that the model they have proposed had given better results when compared with the standard COCOMO model. It was also observed, when provided with enough classification among training data may give better results. But it was also observed that there was non- linearity in the data items.

Jun-qing Li, Quan-ke Pan, Sheng-xianXie,Bao-xian and Yu-ting Wang: at al stated that adaptable employment shop booking issue (FJSP) is critical in numerous examination fields, for example, generation administration and combinatorial improvement. The FJSP issues cover two challenges to be specific machine task issue and operation sequencing issue. In this paper, a mixture of molecule swarm enhancement (PSO) calculation and tabu hunt (TS) calculation are displayed to explain the FJSP with the paradigm to minimize the greatest culmination time (makespan). In the novel cross breed calculation, PSO was utilized to produce a swarm of brilliant competitor arrangements, while TS was utilized to acquire a close ideal arrangement around the given great arrangement. The computational results have demonstrated that the proposed cross breed calculation is proficient and viable for comprehending FJSP, particularly for the issues with expansive scale.

Konstantinos E. Parsopoulos and Michael N. Vrahatis: at el described that the execution of the Particle Swarm Optimization system in adapting to Constrained Optimization issues is explored in this commitment. In the received methodology a non{stationary multi{stage assignment punishment capacity is consolidated, and a few examinations are performed on well{known and broadly utilized benchmark issues. The obtained results are accounted for and contrasted and those got through dierent developmental calculations, for example, Evolution Strategies and Ge- netic Algorithms. Conclusions are inferred and bearings of future reinquiry are uncovered.

Diptam Dutta, Argha Roy, KaustavChaudhury: at el stated that in this paper, the adjustment of system weights utilizing Particle Swarm Optimization (PSO) was proposed as a component to enhance the execution of Artificial Neural Network (ANN) in order of IRIS dataset. Arrangement is a machine learning system used to foresee bunch participation for information examples. To streamline the issue of order neural systems are being presented. This paper concentrates on IRIS plant characterization utilizing Neural Network. The issue concerns the recognizable proof of IRIS plant species on the premise of plant characteristic estimations. Characterization of IRIS information set would be finding examples from inspecting petal and sepal size of the IRIS plant and how the forecast was produced using breaking down the example to structure the class of IRIS plant. By utilizing this example and order, in future forthcoming years the obscure information can be anticipated all the more definitely. Counterfeit neural systems have been effectively connected to issues in example grouping, capacity rough guesses, advancement, and affiliated memories. In this work, Multilayer bolster forward systems are prepared utilizing back engendering learning calculation.

TaghiJavdaniGandomani: et al stated that this article demonstrates a Case Study scrutinize that is performed to look at adequacy of the Planning Poker and Wideband Delphi for two situation studies. Albeit, Wide Band Delphi has been proposed numerous years back and has enough backings in both industry and scholarly zone however, Planning Poker is to some degree new. The study demonstrated that applying these systems in two organizations which were utilizing master's perspective for programming expense estimation, expanded the exactness of expense estimation. Likewise, the outcomes demonstrated that Planning Poker prompted better exactness contrasting with Wideband Delphi. Then again, this exactness is not all that much than Wideband Delphi, it was vital that Planning Poker prompted lessening money related dangers of the undertakings contrasting with Wideband Delphi. The study likewise demonstrated that both Wideband Delphi and Planning Poker diminished the disparages fundamentally.

PoonamRijwani: et al stated that for any industry to stay aggressive, overseeing harmony in the middle of value and expense of programming is critical. Assessing programming improvement exertion remains an unpredictable issue and one which keeps on pulling in impressive exploration consideration. Number of specialists had attempted their endeavors to deliver diverse demonstrating methods in most recent couple of decades. This paper is about the exhaustive distinct investigation of the procedures that were displayed in programming exertion estimation field. In this paper we exhibit the primary discoveries of few examination papers that have used a different parametric and non parametric strategies amalgamated with computational insight procedure, in programming exertion estimation. Every single across the board model talked about at one spot will give scientists a prospect to appreciate the upsides and downsides, likenesses and the distinctions among different models. Decisive words:- programming exertion estimation, algorithmic models, programming Metrics; line of code, computational knowledge, neural systems, back proliferation, mean greatness of relative lapse.

Ali Arshad: et al stated that cost Estimation has constantly remained a core interest zone for analysts of distinctive fields. For programming designing, a superior expense estimation turns out to be more essential with the reality that 66% of every significant venture significantly overwhelm their unique evaluations. Programming expense estimation is a perplexing movement that obliges learning of various key traits about the venture. Then again there are assorted sorts of methods accessible to foresee cost estimation, each of which is having certain upsides and downsides. Line of code (LOC) can be considered as one of the most punctual routines utilized for estimation reason, i.e. how much a lines of code a software engineer/designer has composed on a punch card. With the rise of abnormal state dialects and afterward progression in coding environment, the LOC got to be lacking. Accordingly scientists were compelled to concoct new and creative thoughts in this space. All around the globe, right now numerous devices, systems and techniques are being rehearsed. Pakistan, being one of the vigorously outsourced nation, saying thanks to its shabby work, has likewise utilized a few strategies for expense estimation. This study is centered around Pakistani programming houses with expectation of investigating the product cost estimation strategies, models and methods alongside present condition of hone. This study additionally examination the elements which are influencing the determination of programming expense estimation strategies/models in Pakistan programming industry.

MudasirManzoorKirmani: et al stated that programming task administration is one of the basic exercises in programming advancement process. Numerous product improvement ventures come up short because of poor administration of undertakings. The fundamental point of programming venture administration group is to number what is countable, measure what is quantifiable and get ready quantifiable parameters for consistently screen and deal with a product improvement venture. One of the critical calculates legitimate administration of a product undertaking is programming estimation. At present the vast majority of programming tasks created are utilizing item situated innovation for the improvement of programming activities. Specialists over the globe have utilized article arranged innovation based system known as utilization case point strategy for assessing endeavors for a product venture improvement. This examination work is a push to survey distinctive invariants of utilization case point strategy for exertion estimation. The examination work likewise highlights the significance of persistently refining the UCP strategy in order to reduce the gap between the estimated and actual effort

Mohammad Islam: et al stated that the product business has had noteworthy advance lately. The swwhole existence of programming incorporates two stages: Production and Maintenance. Programming support expense is progressively developing and evaluations demonstrated that around 90%, if programming life expense is identified with its upkeep stage. Extraction and considering the elements influencing programming upkeep expense help to gauge the expense and lessen it by controlling the variables. Cost estimation of upkeep stage is important to anticipate the unwavering quality, enhance the efficiency, undertaking arranging, controlling and flexibility of the product. In spite of the fact that there are

different models to gauge the support expense of customary programming like COCOMO, SLIM, Function Point and so on., yet till now there is no such model to gauge the upkeep expense utilizing fourth era dialect environment. Programming upkeep will keep on existing in the fourth era environment, as frameworks will even now be obliged to advance. In this sort of circumstance there is expected to build up a model to gauge the upkeep expense utilizing fourth era environment. We propose a methodical methodology and advancement for programming upkeep cost estimation model utilizing fourth era dialect environment on the premise of COCOMO II. This model is in view of three parameters: SMCE with Fourth Generation Language Environment, ACT (Annual Change Traffic), Technical and Non-Technical variables which influence the upkeep cost. The ideal results firmly coordinating and it can be accomplished by utilizing model execution.

Ali Idri: et al stated that web Effort Estimation is a procedure of anticipating the endeavors and cost regarding cash, timetable and staff for any product venture framework. Numerous estimation models have been proposed in the course of the most recent three decades and it is accepted that it is an absolute necessity with the end goal of: Budgeting, danger examination, venture arranging and control, and task change speculation investigation. In this paper, we explore the utilization of Fuzzy ID3 choice tree for programming expense estimation, it is planned by coordinating the standards of ID3 choice tree and the fluffy set-theoretic ideas, empowering the model to handle dubious and uncertain information when depicting the product ventures, which can enhance enormously the exactness of got evaluations. MMRE and Pred are utilized, as measures of forecast precision, for this study. A progression of investigations is accounted for utilizing Tukutuku programming undertakings dataset. The outcomes are contrasted and those delivered by three fresh forms of choice trees: ID3, C4.5 and CART.

E. Praynlin, P. Latha: stated that programming Effort Estimation is the procedure of assessing the exertion needed to create programming. By evaluating the exertion, the expense and calendar needed to gauge the product can be decided. Precise Estimate helps the engineer to allot the asset appropriately keeping in mind the end goal to stay away from expense overwhelm and plan invade. A few systems are accessible keeping in mind the end goal to gauge the exertion among which delicate processing based technique assumes a noticeable part. Programming expense estimation manages part of vulnerability among all delicate figuring systems neural system is great in taking care of vulnerability. In this paper Radial Basis Function Network is contrasted and the back spread system and the outcomes are accepted utilizing six information sets and it is found that RBFN is best suitable to gauge the exertion. The Results are accepted utilizing two tests the lapse test and the statistical test.

3. TECHNIQUES (METHODOLOGY)

1. Divide the data items into clusters using k-means clustering.

K-means clustering follows the following steps:

Suppose that we have n sample feature vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ all from the same class, and we know that they fall into k compact clusters, $k < n$. Let \mathbf{m}_i be the mean of the vectors in cluster i . If the clusters are well separated, we can use a minimum-distance classifier to separate them. That is, we can say that \mathbf{x} is in cluster i if $\|\mathbf{x} - \mathbf{m}_i\|$ is the minimum of all the k distances. This suggests the following procedure for finding the k means:

- Make initial guesses for the means $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_k$
 - Until there are no changes in any mean
 - Use the estimated means to classify the samples into clusters
 - For i from 1 to k
 - Replace \mathbf{m}_i with the mean of all of the samples for cluster i
 - end_for
 - end_until
2. Use PSO for parameter tuning of each cluster.

3. Use back propagation algorithm to train the tuned parameters. Algorithm for back propagation follows the following steps:

```
initialize system weights
do
foreach preparing sample ex
forecast = neural-net-output(network, ex)
genuine = teacher-output(ex)
figure slip (expectation - real) at the yield units
figure for all weights from concealed layer to yield layer
figure for all weights from information layer to shrouded layer
upgrade system weights
until all samples grouped accurately or an alternate halting rule fulfilled
give back where its due
4. Compare results.
```

Software detail:

Matlab is a programming environment as well as a high level, interpreted, dynamically typed language. It is well suited for numerical computation, particularly computations involving matrix operations and linear algebra.

4. RESULT

The non linearity of data is removed using clustering technique as well as back propagation neural network algorithm as drivers for the improvement. This method improves the linearity than the MOPOSO model which was used before.

5. CONCLUSION

In our proposed model, we are trying to improve the previous model by removing the non-linearity of the data items which came while using MOPSO model. We are using clustering techniques as well as back propagation neural network algorithm as drivers for the improvement. Results will be compared and the best one will be shown.

ACKNOWLEDGEMENT

Working on this thesis of **Enhancement of multi-objective particles swarm optimization using clustering techniques** provided a unique experience and analysis, I feel great pleasure and privilege in working over this research. I am deeply indebted to “**GGS Kharar**” for the guidance, support and motivation for the many other aids without which it would have been impossible to complete this project. I have no words to express my deep sense of gratitude for Rajan sachdeva (Mentor) mam for her enlightening guidance, directive encouragement, suggestions and constructive criticism for always listening to our problems and helping us out with their full cooperation. Last but not the least, Father J S talyan Mother satyavati talyan Brothers Parbhat talyan and fiancée karan singh who have given me that much strength to keep moving on forward every time, we are greatly thankful to them and have no words to express my gratitude to them.

REFERENCES

- [1] Multi objective Swarm Optimization for cost estimation by G. Sivanageswara Rao, Ch.V. Pani Krishna and K. Rajasekhara Rao, Advances in intelligent systems and computing 248, convention of CSI, volume I
- [2] A hybrid particle swarm optimization and tabu search algorithm for flexible job-shop scheduling problem by Jun-qing Li, Quan-ke Pan, Sheng-xianXie, Bao-xianJia and Yu-ting Wang, International Journal of Computer Theory and Engineering, Vol. 2, No. 2 April, 2010 1793-8201
- [3] Particle Swarm Optimization Method for Constrained Optimization Problems by Konstantinos E. Parsopoulos and Michael N. Vrahatis,

- [4] Training Artificial Neural Network using Particle Swarm Optimization Algorithm by Diptam Dutta, Argha Roy, Kaustav Chaudhury, Volume 4, Issue 5, May 2014 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering.
- [5] A Case Study Research on Software Cost Estimation Using Experts' Estimates, Wideband Delphi, and Planning Poker Technique by Taghi Javdani Gandomani¹, Koh Tieng Wei^{2**} and Abdulelah Khaled Binhamid³
- [6] Software Effort Estimation: A Comparison Based Perspective by Poonam Rijwani¹, Sonal Jain² and Dharmesh Santani³
- [7] A Survey on Software Cost Estimation in the Pakistani Software Industry by Ali Arshad
- [8] Use Case Point Method of Software Effort Estimation: A Review by Mudasir Manzoor Kirmani and Abdul Wahid.
- [9] DEVELOPMENT OF A SOFTWARE MAINTENANCE COST ESTIMATION MODEL: 4TH GL PERSPECTIVE by Mohammad Islam¹, Dr. Vinodani Katiyar
- [10] A Fuzzy Decision Tree to Estimate Development Effort for Web Applications by Ali Idri
- [11] Software Effort Estimation Models Using Radial Basis Function Network by E. Praynlin, P. Latha