Optimization of COCOMO II Coefficients using Cuckoo Optimization Algorithm to Improve The Accuracy of Effort Estimation

I Made Mika Parwita[1], Riyanarto Sarno[2], Alifia Puspaningrum[3]
Department of Informatics Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
[1]mika.parwita14@mhs.if.its.ac.id ,[2]riyanarto@if.its.ac.id, [3]alifia.puspaningrum16@mhs.if.its.ac.id

***Abstract*—Software effort estimation becomes an important factor to support the success of software projects development. Uncertainty and complexity of the software can impact to inaccurate estimation directly; it becomes important challenges to improve the accuracy of estimation. Constructive Cost Model (COCOMO) II used two coefficients to estimate the effort of the software. However, using these coefficients for modern projects often leads to inaccurate estimation. To overcome these problems, Cuckoo Optimization Algorithm is proposed to improve the accuracy of the effort estimation by optimizing the coefficients. Cuckoo Optimization Algorithm (COA) is a new metaheuristic method which has been proven for a great optimization. By utilizing the coefficients of effort estimation using COA, the solution then will be evaluated by using Magnitude of Relative Error (MRE) and Mean Magnitude of Relative Error (MMRE). As the result, the research showed the superiority of our proposed method compared to several methods on Turkish Software Dataset.**

Keywords— COCOMO II Model; Coefficients Optimization; Cuckoo Optimization Algorithm; Software Effort Estimation, Metaheuristic Method

I. INTRODUCTION

In software industries, software effort estimation is a process in estimating effort to develop project of a software system[1]. It is also used to achive an accurate estimation in order to decrease the risk in software development[2]. The important issue in effort estimation is less accuracy of estimation caused by unclear requirements, inconsistency and complexity of software projects [4], [5]. These problems led to the overestimating or underestimating of cost estimation [3]; development team with insufficient budget, lack of labor, and also caused development schedule delay, poor software quality and eventually project failure represented in cost driver of effort estimation model. [6].

Constructive Cost Model (COCOMO) is one of commonly used model for effort estimation. COCOMO II using coefficients to calculate estimation; which play important role in effort estimation [6]. There are two coefficients in this model, i.e. multiplicative and exponent. COCOMO II coefficients considered not suitable to estimate the modern project [4], [10]. These coefficients affect the accuracy of the estimation much. Once the best coefficients is obtained, the optimal effort can be achieved. Therefore, obtaining the optimal coefficients can improve the accuracy of the software estimation. To solve these problems, there were many improved method to estimate the optimal accuracy of effort estimation, such as Nature-inspired metaheuristic techniques.

Various nature-inspired metaheuristic methods have been used in solving many kind of problems, one of them is in software effort estimation. Some of these methods are Simulated Annealing, Genetic Algorithm, Ant Colony Optimization, Particle Swarm Optimization, Cuckoo Optimization Algorithm, and many other. In 2009, Yang and Deb developed Cuckoo Optimization Algorithm (COA) as one of modern metaheuristic methods to solve optimization problems.

In this research, we present an optimization of COCOMO II coefficients in software effort estimation using COA. This method was inspired by Cuckoo Bird behaviour. It's enhanced by Levy Flight which uses a combination of local random walk and the global explorative random walk. It can help the system to not be trapped in a local optimum in obtaining the best solutions [19]. By getting the corresponding best coefficients, software effort estimation will be more accurate [21]. The performance evaluation is analyzed by using Turkish Software Industry dataset.

The structure of this paper is organized as follows: after introduction in Section I, literature review described in Section II. Section III defines several approaches which have been done to improve the accuracy of cost estimation. Section IV presents proposed method which is used in this research. The experiment result presented in Section V. Last section presents the conclusions and future works of this research.

II. LITERATURE REVIEW

This section displays the brief description of COCOMO II and Cuckoo Optimization Algorithm as literature in this research.

A. COCOMO II

Constructive Cost Model (COCOMO) 81 is a model designed to estimate software effort which was published in 1981 by Barry Boehm; the famous scientist who contributed to the development of software project management by creating a scientific approach. However,

978-1-5386-2827-0/17/$31.00 ©2017 IEEE

the complexity and scope of software is growing up, make several limitations for COCOMO [4], [10]. Then, COCOMO II was published in 2000 by Barry Boehm, in order to overcome these problems [5], [12]. Some improvement in several cost drivers from COCOMO 81 was implemented in COCOMO II model. [5], [8], [9].

COCOMO II includes several attributes, i.e. 17 Effort Multipliers (EMs), 5 Scale Factor (SFs), and Software Size (SS) [5], [8], [9]. The 17 Ems are described in Product, Personnel, Computer, and Project categories. Table I shows 17 EMs of COCOMO II and SFs displayed in Table II.

$$\text{Effort (PM)} = \text{A x (size)}^{\text{E}} \text{ x } \prod_{i=1}^{12} EM_{\text{i}} \quad (1)$$

$$\text{E} = \text{B} + (0.01 \text{ x } \sum_{i=1}^{n} SF_{\text{i}}) \quad (2)$$

Equation (1) is used to estimate effort of the software. Effort is estimated in Person-Month (PM) which obtained with multiplicative coefficient (A) multiplied by the size of the software (size) in units Kilo-Source Lines of Code (KSLOC) power E, then multiplied by multiplications of Effort Multiplier (EM) as much as *i* data. E value obtained by (2), i.e. the sum of Scale Factor (SF) as much as *i* data then multiplied 0,01 then added by exponent coefficient (B). The A coefficient is determined by the value of 2.94 and B has a value of 0.91; where this value has been determined based on historical data by the data 161 points [8].

TABLE I. EFFORT MULTIPLIERS OF COCOMO II

Category	Effort Multipliers	Description
Product Attributes	RELY	Required Software Reliability
	DATA	Database Size
	CPLX	Product Complexity
	RUSE	Developed for Reusability
	DOCU	Documentation Match to Life Cycle Needs
Computer Attributes	TIME	Execution Time Constraint
	STOR	Main Storage Constraint
	PVOL	Platform Volatility
Personnel Attributes	ACAP	Analyst Capability
	PCAP	Programmer Capability
	PCON	Personnel Continuity
	APEX	Application Experience
	PLEX	Platform Experience
	LTEX	Language and Tool Experience
Project Attributes	TOOL	Use of Software Tools
	SITE	Multisite Development
	SCED	Required Development Schedule

TABLE II. SCALE FACTORS OF COCOMO II

Scale Factors	Description
PREC	Precedentedness
FLEX	Development Flexibility
RESL	Risk Resolution
TEAM	Team Cohesion
PMAT	Process Maturity

B. Cuckoo Optimization Algorithm (COA)

Cuckoo Optimization Algorithm (COA) was proposed by Yang and Deb in 2009 [13]. This algorithm was developed to solve optimization problems and proved to be faster and simpler in finding the best solution than other metaheuristic methods, e.g. Genetic Algorithm (GA), Firefly Algorithm (FA), Particle Swarm Optimization (PSO), and Artificial Bee Colony (ABC) [14], [15].

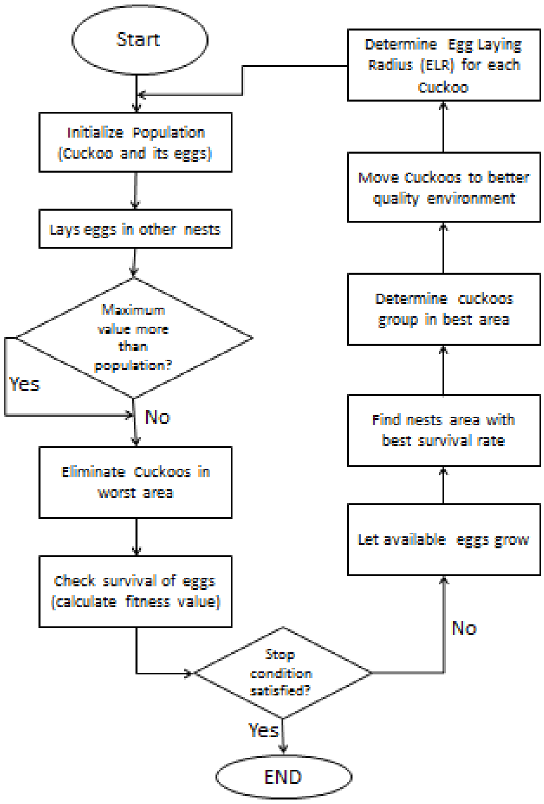

Fig. 1. Flowchart of Cuckoo Optimization Algorithm

This algorithm is derived from behaviour of the Cuckoo bird species; lays their eggs in other bird's nests, which is not species of Cuckoo. If the host nest realized there are other eggs that not belong to her, there are two possibilities: removes the eggs or leave the nest and then make a new nest.

Eggs in a nest representing the solutions are to be sought [13], [16]. A nest is considered only have one Cuckoo egg; the purpose of this step is to replace the poor eggs with the eggs which has better quality as solution. In the nest, each egg is considered as a unique solution, different from any other solutions. After Cuckoos successfully grown up, mature Cuckoos will look for other host nests for laying their eggs for the next generation. The process is repeated until all Cuckoos gathered in one nest. Thus, the COA may be used to find optimal solutions of the problems that can be formulated mathematically in terms of optimization. COA is based on three basic principles [14]:

- Each Cuckoo lays its eggs in random host nests.
- Nest with high quality eggs (solutions) are considered as the best nest for next generation.
- Number of host is fixed and each nest can discover eggs with probability pa [0.1].

Fig. 1 shows the flowchart of the COA. The algorithm started by initializing the population - the Cuckoo birds. Then lay eggs in other host nests as a random solution which will represent the next generation. Cuckoo in worst area will be eliminated if number of population generated more than maximum value. The laying of the eggs has maximum radius for each of the Cuckoo.

$$ELR_i = r \times \frac{SumEgg_i}{TotalSumEgg} \times (posMax - posMin) \quad (3)$$

The spread of eggs is determined by *Egg Laying Radius* (ELR) [16]. Equation (3) describes the ELR of the i^{th} Cuckoo. The maximum distance of i^{th} Cuckoo (ELRi) was obtained using a variable radius (r) as the maximum distance Cuckoo, the highest position limit (posMax), lowest position limit (posMin), and the number of eggs. By using ELR, determining the value of the random laying eggs more controlled but not reduce the performance of the algorithm [15], [16]. Then, evaluate the value of fitness for each solution to find the best solution.

III. Related Work

Various optimization techniques have been used to optimize the COCOMO II coefficients. Genetic Algorithm (GA), Tabu Search, Particle Swarm Optimization (PSO), and Harmony Search are recently algorithms used to optimizing coefficients in effort equation. In [2] has utilized GA for optimized COCOMO coefficients. An evolutionary model developed by GA to estimate the effort. The result of the model is that new model successfully improves the performance by reducing MMRE. In [17], the combination of GA and Tabu Search used to obtain the values of the COCOMO coefficients. The datasets of NASA are used as input data sets. The results of experiments based on Mean Absolute Relative Error (MARE), the hybrid of GA and Tabu Search is successfully improves the accuracy compared to COCOMO. Another technique that implemented is PSO. In [11], PSO algorithm used to evaluate the COCOMO model coefficients. The COCOMO model is used as the comparison and coefficient that produced by PSO decreased MMRE than COCOMO model. Furthermore, in [20], PSO can optimized coefficients and reduce the uncertainties. Harmony Search implemented to obtain the better coefficient in [6]. Harmony Search and COCOMO produce coefficients that tested with NASA dataset. As the results, proposed method decrease MMRE in organic, semidetached, and embedded COCOMO.

Combination of Fuzzy Logic and Local Calibration also implemented to increase the accuracy. In [18], Fuzzy Logic used to recalculate Effort Multiplier and Local Calibration implemented to calibrate the suitable coefficient COCOMO II model. The data sets are collected from Turkish Software Industry. These datasets are used to calibrate COCOMO II coefficients. As the results, Mean Magnitude of Relative Error (MMRE) value as one of evaluation compared the performance of proposed method and COCOMO II.

From the discussion of the literature, it is observed that most of the reported methods are trying to obtain the best value of coefficient so that MMRE value can be reduced. Cuckoo Search Algorithm as one of modern metaheuristic algorithm proposed a technique to obtain the optimal solution which superior in terms of covergence characteristic and less in number of parameter which is tuned [19]. Hence, implementing Cuckoo Optimization Algorithm to solve various domain is still growing unceasingly .

IV. Methodology

One of the major problems in software development process is to estimate the cost. Several estimation models have been proposed but only few can reach the level of satisfaction. Among all of the existing estimation methods, COCOMO is the most popular. But due to increase of complexity and overdemanding software requirements it becomes less accurate. The main objective of this research is to immprove the accuracy of COCOMO II effort estimation with optimizing the coefficients in (1) and (2) by using COA and MRE used for checking the performance. For better performance results should be minimum MRE and MMRE. In application of this method, dataset Turkish Software Industry is used as input data. Table III displays the parameters of COA. These are the steps of the proposed method:

Input : Dataset Turkish Software Industry (Effort Multiplier, Scale Factor, Actual Effort, and Size), Parameter COA, shown in Table III
Output : MMRE, Coefficient of A and B

Table III. PARAMETERS OF COA

Parameter	**Value**
Number of cuckoo	5
Number of eggs	Minimal = 2 Maximal = 4
Iteration	500
Dimension	2
Position	Minimal = -10 Maximal = 10
Minimum function	0
Radius	10
Maximum number of cuckoo	10
Limit position difference eggs	0.000000000001
Number of cluster	1
Lambda	9

The procedure for implementing the method is explained well in the following steps:

Step 1: Initialization population of Cuckoo and count the number of egg laying central point of each Cuckoo.

Step 2: Specify the number of eggs in randomly for each Cuckoo.

Step 3: Calculate the radius of laying eggs (ELR) from each new eggs using (3).

Step 4: Calculate position of new eggs for each Cuckoo.

Step 5: Checking the number of existing Cuckoo not to exceed the maximum number of Cuckoo.

Step 6: Calculate fitness value for each Cuckoo using COCOMO II effort estimation equation in (1) MRE, and MMRE

Step 7: Find the best MMRE and specify the center point of the best Cuckoo

Step 8: Remove Cuckoo that exceeded the maximum number of Cuckoo

Step 9: Perform data classification on the best Cuckoo and specify the center point

Step 10: Add Cuckoo with random positions around the best Cuckoo position

Step 11: If the stop criteria is satisfied, stop the algorithm and display best MMRE and the coefficients (A and B). If not, repeat from Step 2

V. Evaluation Criteria and Data Set

A fundamental question that needs to be asked in all estimation methods is how accurate are the estimation. This research tries to compare estimated effort with actual effort as the evaluation using most common criteria as fiteness function on software effort estimation, namely: Magnitude of Relative Error (MRE) and Mean Magnitude of Relative Error (MMRE).

The MRE is used to determine the accuracy between actual with predicted effort. MRE is calculated for each project as defined in Equation (4):

$$\text{MRE} = \frac{|\,ActualEffort - EstimatedEffort\,|}{ActualEffort} \text{ x } 100 \qquad (4)$$

After calculating the effort for each project, number of accuracy for each project is averaged using Mean-MRE (MMRE) as defined in Equation (5):

$$\text{MMRE} = \frac{1}{N}\sum_{i=1}^{N} MRE_{\text{i}} \qquad (5)$$

The better accuracy will be measured based on the value of MMRE. The smaller MMRE indicates the better accuracy which means that the proposed method can estimate the effort close to the actual value. Performance of the proposed method's also compared to another existing methods in estimating effort such as Fuzzy Local Calibration and COCOMO II.

VI. Experimentations And Result

Proposed method is assessed on Turkish Software Industry dataset with 12 data points were used for experiment. Each data point has consisted of 25 attributes project id, 5 scale factor, 17 effort multipliers in the range of very low to extra high, software size in SLOC, and actual effort.

This data will be used to obtain the optimal coefficient multiplicative (A) and exponential (B). The results of optimization can be used to estimate new projects of the similar category. As the result of optimization experiment, the coefficient multiplicative (A) is changed from 2,94 and 4,4174. While the value of the exponent (B) is changed from 0,91 becomes -0,1847.

Table IV. Actual Effort and Estimated Effort

Project Id	Actual Effort	COCOMO II	Fuzzy Local Calibration	COA
1	1.2	3.49	1.74	1,57
2	2	2.86	1.89	2,01
3	4.5	9.30	3.57	2,87
4	3	33.98	7.30	4,10
5	4	63.16	10.65	4,90
6	22	27.73	2.51	0,72
7	2	2.29	0.98	0,74
8	5	147.09	14.32	5,00
9	18	297.60	12.84	2,50
10	4	64.00	7.99	3,09
11	1	0.92	0.81	0,98
12	2.1	2.04	1.53	1,82

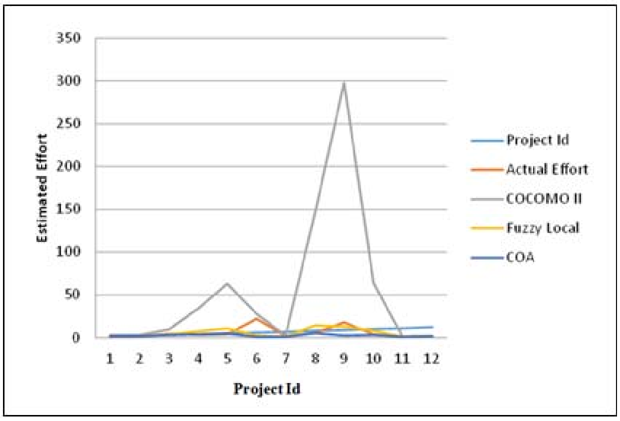

Fig.2. Comparison actual effort to estimated effort

Table IV shows the comparison among the actual effort and estimated effort which obtained using other models. It can be observed that the estimation obtained by COA is closer to the actual effort, compared to conventional COCOMO II and Fuzzy Local Calibration. The graphical representation of Table IV is shown in Fig. 2. According to this figure, estimated effort of COA is almost close to actual effort in 8 points from 12 points. Using this comparison, we can measure the improvement by using the deviation between actual and estimated effort, which represents how well the method in estimating effort. In addition, it has to be considered that small number of deviation between actual and estimated effort can be one of indication to obtain better accuracy.

TABLE V. COMPARISON ACCURACY IN MRE AND MMRE

Project Id	MRE		
	COCOMO II	*Fuzzy Local Calibration*	*COA*
1	190.67	44.98	31,20
2	42.84	5.41	0,49
3	106.76	20.64	36,26
4	1032.58	143.22	36,83
5	1478.89	166.35	22,38
6	26.05	88.61	96,71
7	14.43	50.87	62,81
8	2841.79	186.47	0,01
9	1553.36	28.65	86,12
10	1499.90	99.79	22,72
11	7.61	18.92	0,01
12	2.75	27.05	0,13
MMRE	**733.14**	**73.41**	**34.20**

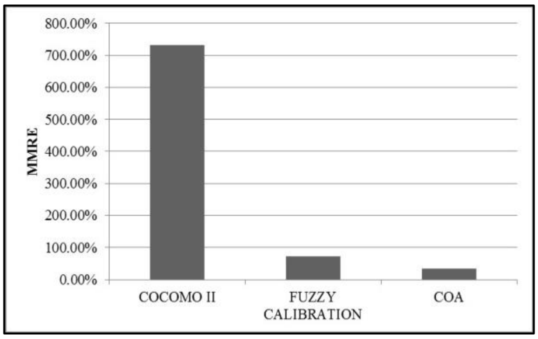

Fig. 3. Comparison of MMRE

Table V displays the experiment result in MRE and MMRE. The proposed method can show better solution in almost project than the comparison methods, which for each data point, the new coefficients produces best MRE value in eight data point, i.e. 1, 2, 4, 5, 8, 10, 11 and 12. On the other hand, Fuzzy and Local Calibration method produces the best MRE in two data points, i.e. 3 and 9 because the effort multiplier and scale factor value reached a small margin with actual effort. In the other hand, conventional COCOMO II produces the best MRE just in two data points, i.e. 6 and 7. It can be observed, from the 12 data points, 66,67% of the MRE project is obtained by using optimized coefficients. It implies that the proposed method performs well in estimating effort compared to comparison methods with respect to the objective function.

Fig. 3 briefly observed that the MMRE of COA is the smallest value compared to COCOMO II and Fuzzy Calibration. It shows that COA is more accurate in estimation effort. The new coefficients obtained through optimization using COA produce MMRE of 34.2. This value represents the best MMRE value compared to conventional COCOMO II which produces 733.14 and Fuzzy Local Calibration produces 73.41. This numbers represent the new coefficients reduce MMRE 698.94% compared to conventional COCOMO II and reduce 39.91% compared to MMRE produced by Fuzzy Local Calibration. The MMRE (%) indicates that, for general dataset, the proposed COA method can obtain better optimal solution than those by other approaches. By using levy flight, the proposed method can find the optimal solution which can generates lower MMRE. So that, the proposed method can reduce MMRE 698.94% compared to conventional COCOMO II. In addition, it emphasizes that as close as the predicted effort to the actual effort mostly generate smaller value of MRE, and as a result the minimum value of MMRE will be obtained.

For the better performance of Cuckoo Optimization Algorithm in effort estimation, a hybrid method can be optimalized to get better initialization and also exploit the useful information to generate better quality solutions. In addition, applying to the newest dataset can becomes further research in order to prove the feasibility of the algorithm to optimize the coefficients.

VII. CONCLUSION

COCOMO II coefficients become important factor determinant of accuracy of the estimation. In this paper, COA is implemented to optimize the multiplicative (A) and exponent (B) coefficients of effort estimation equations to improve the accuracy of COCOMO II. According to the experiment result, the new coefficients provide higher accuracy than coefficients of conventional COCOMO II and Fuzzy Local Calibration. The new coefficients reduce MMRE up to 698,94% compared to COCOMO II and reduce 39,91% compared to Fuzzy Local Calibration.

REFERENCES

[1] B. K. Singh and K. A. Misra, “Software Effort Estimation by Genetic Algorithm Tuned Parameters of Modified Constructive Cost Model for NASA Software Projects,” *International Journal of Computer Applications*, vol. 59, no. 9, pp. 22–26, 2012.

[2] S. Bilgaiyan, S. Mishra, and M. Das, “A Review of Software Cost Estimation in Agile Software Development Using Soft Computing Techniques,” *2016 2nd International Conference on Computational Intelligence and Networks (CINE)*, pp. 112–117, Jan. 2016.

[3] R. Sarno, J. Sidabutar, and Sarwosri, “Comparison of different Neural Network architectures for software cost estimation,” *2015 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, pp. 68–73, Oct. 2015. doi:10.1109/IC3INA.2015.7377748

[4] H. Rastogi, S. Dhankhar, and M. Kakkar, “A Survey on Software Effort Estimation Techniques,” *2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, pp. 826–830, 2014.

[5] B. Boehm, B. Clark, E. Horowitz, and C. Westland, “Cost Models for Future Software Life Cycle Processes: COCOMO 2.0,” *Annals of Software Engineering*, vol. 1, no. 1, pp. 57–94, 1995.

[6] S. M. S. Jafari and F. Ziaaddini, “Optimization of Software Cost Estimation using Harmony Search

Algorithm," *1st Conference on Swarm Intelligence and Evolutionary Computation (CSIEC2016)*, pp. 131–135, 2016

[7] B. W. Boehm, "Software Engineering Economics," *IEEE Transactions on Software Engineering*, vol. SE-10, no. 1, pp. 4–21, Jan. 1984.

[8] B. Clark, S. Devnani-Chulani, and B. Boehm, "Calibrating the COCOMO II Post-Architecture model," *Proceedings of the 20th International Conference on Software Engineering*, pp. 477–480, 1998

[9] R. Sarno, J. Sidabutar, and Sarwosri, "Improving the Accuracy of COCOMO's Effort Estimation Based on Neural Networks and Fuzzy Logic Model," *2015 International Conference on Information, Communication Technology and System (ICTS)*, pp. 197–202, 2015. doi:10.1109/ICTS.2015.7379898

[10] P. K. Patil, "A Review on Calibration Factors in Empirical Software Cost Estimation (SCE) Models," *International Journal of Software Engineering Research & Practices*, vol. 3, no. 2, pp. 1–7, 2013.

[11] F. S. Gharehchopogh, I. Maleki, and S. Reza Khaze, "A novel particle swarm optimization approach for software effort estimation," *International Journal of Academic Research*, vol. 6, no. 2, pp. 69–76, Mar. 2014

[12] J. Kaur and R. Sindhu, "Parameter Estimation of COCOMO II using Tabu Search", vol. 5, no. 3, pp. 4463 – 4465, 2014.

[13] X. S. Yang and S. Deb, "*Cuckoo Search via Levy Flights",* Conference Proceedings of World Congress on Nature & Biologically Inspired Computing, New York, IEEE Publisher, pp. 210 – 214. 2009

[14] X. S. Yang and S. Deb, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4, pp. 330–343, 2010.

[15] E. Shadkam and M. Bijari, "Evaluation the Efficiency of Cuckoo Optimization Algorithm," *International Journal on Computational Science & Applications*, vol. 4, no. 2, pp. 39–47, Apr. 2014.

[16] R. Rajabioun, "Cuckoo Optimization Algorithm," *Applied Soft Computing*, vol. 11, no. 8, pp. 5508–5518, Dec. 2011.

[17] F. S. Gharehchopogh, R. Rezaii, and B. Arasteh, "A new approach by using Tabu search and genetic algorithms in Software Cost estimation," *2015 9th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 113–117, Oct. 2015.

[18] M. Baiquni and R. Sarno, "Optimasi pada Constructive Cost Model (COCOMO) untuk Meningkatkan Akurasi," *Jurnal Teknik ITS*, vol. 4, no. 1, 2015.

[19] Yang, X-S., & Deb, S. "Cuckoo search and Firefly Algorithm Theory and Application", pp. 1 – 12, 2013.

[20] Parkash Jai and Karambir , "COCOMO II Model Paramter Optimization using PSO and Effort Estimation", *International Journal of Information Technology and Mechanical Engineering (IJITME)*, vol. 1, pg. 1-11, Jul. 2014

[21] R.R. Putri, R. Sarno, D. Siahaan, A. S. Ahmadiyah, and S. Rochimah, "Accuracy Improvement of The Estimatios Effort in COCOMO II Based on Logic Model of Fuzzy", *Advanced Science Letters,* vol.23, no.3, pp.2478-2480, March 2017. doi:10.1166/asl.2017.8767