



4th Information Systems International Conference 2017, ISICO 2017, 6-8 November 2017, Bali, Indonesia

A Hybrid Cuckoo Optimization and Harmony Search Algorithm for Software Cost Estimation

Alifia Puspaningrum, Riyanarto Sarno*

Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Abstract

Nowadays, software effort and time estimation becomes a substantial factor in software engineering community. Accurate estimation can enhance performance in managing the project schedule, human resource allocation, cost estimation, etc. Constructive Cost Model (COCOMO) II is commonly used model of software estimation influenced by four coefficients. The optimal coefficients will produce optimal model in estimating effort and time development. To overcome that problem, this paper proposed a hybrid model of cuckoo search and harmony search algorithm for optimizing four coefficients of COCOMO-II to get optimal estimation. The proposed approach is applied on NASA 93 dataset and evaluated by using Magnitude of Relative Error (MRE) and Magnitude of Relative Error (MMRE). Experimental results show that the proposed method is more effective in estimating effort and time development of the software project than COCOMO-II and cuckoo search algorithm.

© 2018 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 4th Information Systems International Conference 2017.

Keywords: COCOMO II Model; Effort Estimation; Time Development Estimation; Cuckoo Search; Harmony Search; Hybrid Optimization Parameter

1. Introduction

Software effort and time development estimation models are hot topic of study over three decades for software engineering community¹. Good estimation can enhance performance of the company especially in managing the project schedule, human resource allocation, cost estimation, etc². Those advantages can reduce the failure

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .

E-mail address: riyanarto@if.its.ac.id

probability or project delay³. So that, estimating software effort and time development is one of the most crucial responsibility for all the people involved with software project management.

There are some models to estimate value of software project's effort and time development of software project, one commonly used is Constructive Cost Model (COCOMO) II. It consists of coefficients which affect the accuracy of the estimation much. In estimating four parameters of COCOMO II, there are many method to estimate the optimal value, such as machine learning method. Currently, the modelling capabilities of effort and its components in software development makes machine learning methods attract more attention⁴. The aim of machine learning method in estimating effort and time development is find the optimal solution. Some of machine learning methods are include Ant Bee Colony Algorithm⁵, Differential Evolution⁶, Firefly Algorithm², Genetic Algorithm⁷, Harmony Search⁸, Cuckoo Search⁹, Neural Network^{1,10,11}, Fuzzy Logic Model¹², etc. Furthermore, the improved efficiency is often achieved through the hybridization of the algorithm and deterministic optimization techniques. Improving the searching ability of the optimal solutiois is the aim of algorithm hybridization. For example, hybrid model of artificial bee colony and genetic algorithms¹³, hybrid model of cuckoo optimization and K-Nearest Neighbors¹⁴, hybrid model of firefly and genetic algorithm¹⁵, hybrid model neural networks and fuzzy logic model¹⁰, hybrid model bee colony and chaos optimization algorithm¹³, etc.

In 2009, Yang and Deb developed cuckoo search (CS) inspired by cuckoo species lifestyle. In addition, CS can give optimum solution but the searching process using levy flight can not be guaranted. In order to overcome the problem, Harmony Search (HS) can be one of tecnique to be combined with cuckoo search. HS can provide mutation operator to the Cuckoo Search method. So that, the exploitation capability of the solution will increase.

By using the advantages of CS and HS, this paper proposes a hybrid model to estimate the optimal coefficients of software effort and software time development based on cuckoo search and harmony search (CSHS) algorithm by using NASA 93 dataset. This study focuses on the application of the CS algorithm to the software effort and time estimation design problems in the multi-objective domain, and more importantly, attempts to improve its performance through the incorporation of the HS mechanism.

The research paper is organized as follows: Section 2 contains a description of COCOMO II, Cuckoo Search, and Harmony Search. Section 3 explains about the proposed method of hybrid model cuckoo search and harmony search algorithm. Section 4 gives various evaluation criteria. Section 5 explains experimental results. Finally, Section 6 concludes the proposed method and the future work.

2. Background

2.1. Constructive Cost Model (COCOMO) II

COCOMO II is a cost estimation model developed by Barry W. Boehm in 1981 for estimating effort and time development by collecting data from vast application projects¹⁰. COCOMO II consist of several parameters such as 17 effort multipliers, 5 scale factors, and software size. By using the parameters, we can estimate the value of Person Month (Effort) and Time Development for each project. Person-Months and Time Development is influenced by four coefficients of COCOMO-II, namely A, B, C, and D which has been declared before.

1. Effort Estimation is declared as Person-Months (PM) in COCOMO-II. PM is stated as how long software defelopment project will be worked by one person in one month. PM can be calculated by using Equation (1).

$$PM = A \times Size^E \times \prod_{i=1}^{17} EM_i \quad (1)$$

where A = 2,94 Size is described as the project line of code. In additon, EM is effort multipliers that has been presented in Table 2.

2. Five scale factors (SF) aggregation declared in Table 1 influences the value of Exponent E of Equation (1). Exponent E can be calculated by using Equation (2).

$$E = B + 0.01 \times \sum_{j=1}^5 SF_j \quad (2)$$

where $B = 0,91$

3. Time Development (TDEV) Estimation or sometimes stated as schedule estimation is the estimation how long the project will be developed in months. TDEV can be calculated by using Equation (3).

$$\text{TDEV} = C \times (PM)^F \quad (3)$$

where $C = 3.67$

4. Equation (4) will be used to calculate the value of Exponent F in Equation (2)

$$F = D + 0.2 \times 0.01 \times \sum_{j=1}^5 SF_j \quad (4)$$

where $D = 0.28$

The coefficient value of A, B, C, D then will be toned by using hybrid model of cuckoo search and harmony search for NASA 93 dataset.

2.2. Cuckoo Search (CS)

In 2009, Yang and Deb developed Cuckoo Search Algorithm as one of optimization algorithm inspired by nature¹⁶. CS is inspired by cuckoo lifestyle and its characteristics in egg laying in other host birds. In this algorithm, each egg in a nest is defined as solution and a new solution is defined as cuckoo egg. The idea is replacing a previous solution which is not so good with the new and better solution in the nest, following the rules below¹⁷:

1. In initialization, each cuckoo will lay one egg in a chosen host nest randomly
2. The best nest (solutions) will be carried over to the next generation
3. Each host bird can throw egg that is not their egg or abandon the nest with probability p_a [0,1], then find a new location to build a new nest.

2.3. Harmony Search (HS)

Harmony search is one of meta-heuristic developed by Geem in 2001¹⁸. The principles of musical harmony improvisation is the idea of HS¹⁸. This algorithm does not need any information such as gradient for the objective function. So that, the algorithm provide its simplicity as one of its advantages. There are three parameters using in the optimization process, namely harmony memory consideration rate (HMCR), pitch adjustment rate (PAR), and distance bandwidth(bw). In addition, harmony memory size, the number of improvisations(NI) also include as parameters.

3. Proposed Method

To achieve better accuracy, the proposed method combines the optimization capabilities of Cuckoo Search and Harmony Search. In addition, it is used to improve the exploitation capability of the proposed method so as to not being trapped into local optima. The proposed hybrid method of cuckoo search and harmony search has been proposed and explored by many studies, for example optimal design of water distribution systems¹⁹, cancer classification²⁰, knapsack problems²¹, global numerical optimization²², etc.

In estimating software effort and software time development, the proposed method uses double stage of optimization technique, cuckoo search in the first stage and harmony search in the second one. Cuckoo search provides local search and global search in finding the optimal solution. Therefore, harmony memory and pitch adjustment rate of Harmony Search algorithm can ensures that the local solutions are retained while the randomization makes the proposed method able to globally explore the solution space more effectively.

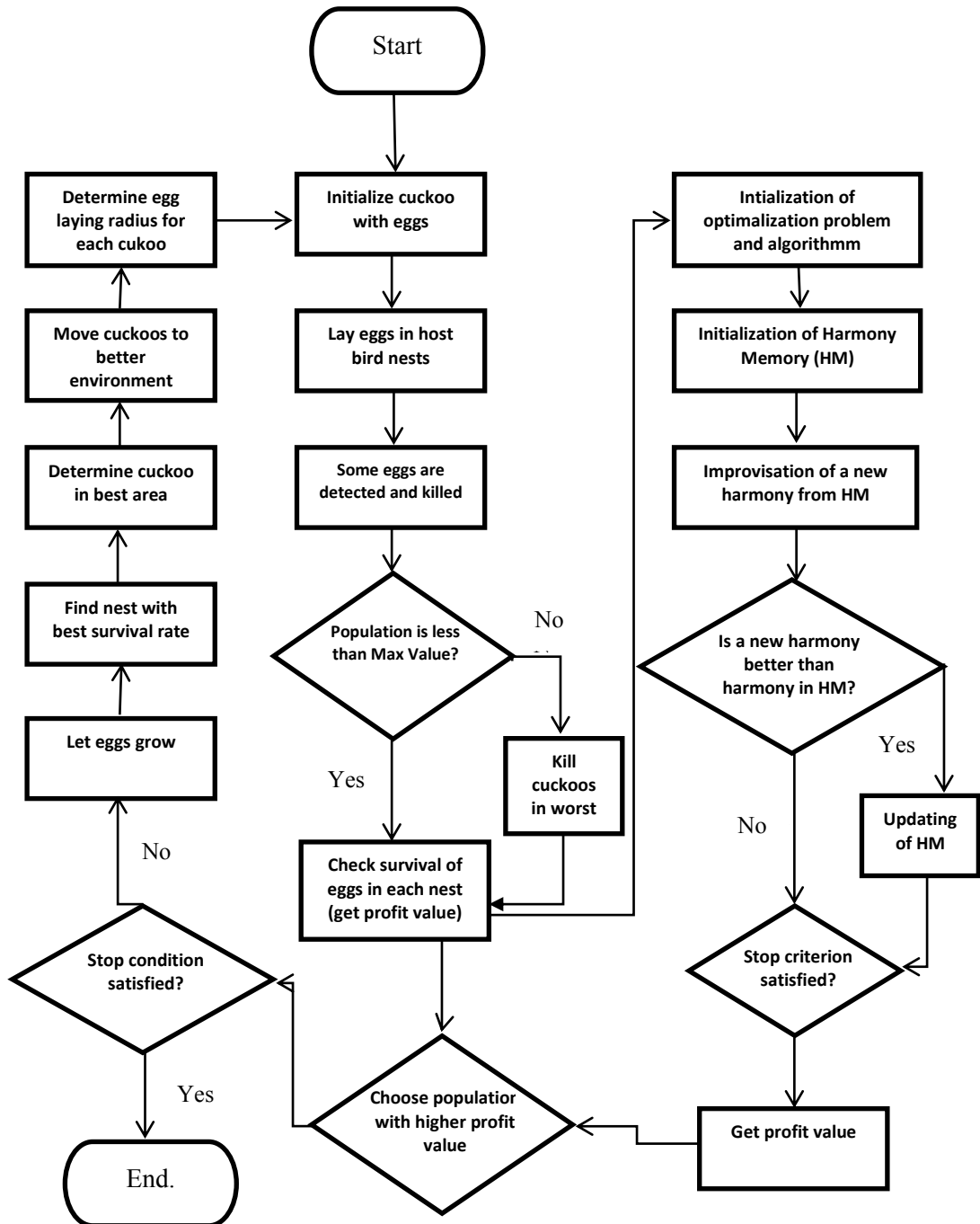


Fig. 1. Flowchart of CSHS algorithm

CSHS begins with random generation of the initial population (or N host nests). After finding the best solution (X_{global}) in the first stage, a new harmony vector (solution) is created based on the HS strategy. After that, the harmony vector value will be compared by X_{global} . Then, HM_{best} vector substitutes the X_{global} only if it has better fitness. This procedure is reiterated as many as the number of iteration.

For software effort and time development estimation, the description of implemented steps is described in this following steps:

- Each of the N populations represents a possible combination of four coefficients, namely A, B, C, and D.
- Compute the value of person-months by using A and B coefficients.
- The value of person-months, C, and D coefficients will be used to calculate the value of time development.
- Compute the subtraction between actual time development and estimated value.
- The procedure a-d will be repeated as many as the number of solution
- The average value of subtraction is used as the fitness value for each of the solution.
- The smallest value of fitness will be declared as the best solution.

4. Evaluation Method

The proposed method will be measured by using Magnitude of Relative Error (MRE) and Mean Magnitude of Relative Error (MMRE). In addition, MMRE will be defined as fitness function for the hybrid model. MRE is measured by using Equation 5:

$$\text{MRE} = \frac{|Effort_i - Effort_i|}{Effort_i} \quad (5)$$

Equation (6) will be used to calculate Mean Magnitude of Relative Error (MMRE) .

$$\text{MMRE} = \frac{1}{n} \sum_{i=1}^n \frac{|Effort_i - Effort_i|}{Effort_i} \quad (6)$$

The lower value of MMRE will be stated as better performance. The evaluation method will be used to evaluate the propose method performance in estimating the value of effort and time development compared to COCOMO II coefficients value.

5. Experiments

The experiment applied a hybrid model of cuckoo search and harmony search for optimizing the value of COCOMO II parameters, namely A, B, C, and D. The parameters will be used to measure the value of effort and time development estimation. The goal is obtaining closer value of the estimation to the actual value. The proposed method will be assessed by using NASA 93 which consists of 93 projects.

Table 1. Performance of Effort Estimation and Time Development MMRE.

Project No.	Effort Estimation MRE			Time Development MRE		
	COCOMO	CS	CSHS	COCOMO	CS	CSHS
1	38,38	28,62	29,90	0,26	0,09	0,29
2	41,58	32,12	33,20	0,25	0,09	0,29
3	33,94	17,69	15,15	0,09	0,00	0,18
4	38,89	24,15	22,00	0,10	0,01	0,18
5	3,91	27,68	30,41	0,12	0,00	0,19
6	33,02	10,05	2,50	0,03	0,02	0,10
7	15,71	10,10	17,14	0,05	0,03	0,12
8	45,33	40,16	43,41	0,57	0,35	0,41
9	50,62	38,38	36,41	0,36	0,16	0,08
10	65,42	59,31	59,63	0,65	0,53	0,29
11	64,08	54,58	52,71	0,27	0,26	0,08
12	62,39	59,83	62,62	2,15	1,34	1,18
13	38,96	25,69	24,56	0,42	0,32	0,09
14	49,81	60,03	48,89	2,50	0,68	0,85
15	36,27	25,02	25,60	0,69	0,53	0,28

Project No.	Effort Estimation MRE			Time Development MRE		
	COCOMO	CS	CSHS	COCOMO	CS	CSHS
16	21,37	16,01	21,86	1,98	0,43	0,52
17	0,13	4,12	4,68	3,47	1,58	1,81
18	8,75	4,47	1,80	0,88	0,57	0,31
19	52,91	43,63	43,42	0,54	0,44	0,20
20	71,06	95,48	90,24	1,14	0,51	0,28
21	9,04	7,11	6,34	0,20	0,06	0,27
22	35,70	29,62	33,45	0,57	0,35	0,41
23	30,89	20,56	22,39	0,30	0,11	0,31
24	45,77	35,09	34,85	0,46	0,19	0,06
25	35,98	27,52	29,91	0,99	0,22	0,04
26	48,01	36,24	34,95	0,31	0,19	0,04
27	30,72	17,20	16,97	0,36	0,11	0,13
28	33,36	25,66	28,80	0,89	0,04	0,09
29	37,27	25,29	25,25	0,38	0,11	0,12
30	35,42	21,96	21,17	0,31	0,10	0,13
31	37,93	29,09	31,00	0,65	0,08	0,12
32	39,97	31,76	33,83	0,71	0,09	0,10
33	19,20	2,71	7,32	0,07	0,02	0,15
34	43,71	31,12	29,84	0,12	0,01	0,21
35	36,16	23,27	22,77	0,15	0,02	0,23
36	36,18	19,68	16,64	0,23	0,08	0,14
37	19,45	2,76	1,83	0,39	0,12	0,12
38	47,43	43,49	47,20	1,95	0,67	0,67
39	30,96	14,18	11,67	0,27	0,11	0,12
40	41,45	30,19	30,11	0,49	0,10	0,14
41	42,14	40,30	45,72	1,30	1,67	0,84
42	59,97	59,99	64,38	3,28	0,84	0,59
43	51,88	51,70	56,89	3,99	0,80	1,73
44	19,69	19,43	28,11	3,83	0,21	1,36
45	50,98	46,89	50,12	2,21	1,03	0,96
46	67,86	68,51	72,35	5,36	1,39	2,75
47	37,89	36,16	42,11	3,10	1,66	1,76
48	48,33	42,28	44,68	1,35	0,60	0,42
49	27,83	49,95	48,50	0,92	0,06	0,15
50	29,26	23,30	27,92	2,04	0,20	0,10
51	47,50	36,12	35,17	0,46	0,14	0,12
52	41,29	30,78	31,22	0,68	0,12	0,13
53	41,59	37,64	42,00	2,59	0,14	0,54
54	49,36	48,39	53,46	4,50	0,77	1,06
55	40,06	33,25	36,15	1,44	0,13	0,08
56	34,76	33,66	40,27	1,99	1,26	0,28
57	35,11	29,19	33,17	0,46	0,37	0,45
58	73,48	62,42	57,78	0,01	0,02	0,06
59	249,98	226,00	176,85	20,79	26,48	2,44
60	25,83	26,51	34,97	3,65	1,00	1,72
61	49,02	44,37	47,50	2,23	0,87	0,85
62	61,40	61,16	65,27	5,92	1,25	2,84
63	3,79	11,57	4,24	1,79	0,83	0,73
64	52,24	46,10	47,98	0,98	0,57	0,33
65	45,11	42,46	47,14	3,18	1,19	1,45
66	28,63	25,59	31,89	3,41	0,15	0,97
67	148,31	146,51	118,42	6,45	1,04	3,06
68	162,43	165,99	138,96	4,49	1,72	2,38
69	1,71	6,30	2,53	0,60	1,94	1,34
70	1,41	5,68	3,29	0,47	1,43	1,08

Project No.	Effort Estimation MRE			Time Development MRE		
	COCOMO	CS	CSHS	COCOMO	CS	CSHS
71	42,42	62,31	57,67	0,11	0,39	0,56
72	12,60	20,42	12,13	0,46	1,10	0,93
73	3,27	4,33	2,29	0,40	0,96	0,87
74	78,92	75,19	75,39	0,04	0,19	0,39
75	47,44	44,20	48,30	0,34	1,07	0,93
76	1,94	5,79	3,45	1,07	2,28	1,32
77	54,79	53,32	35,65	1,98	5,05	2,01
78	689,59	718,47	646,41	0,77	2,47	1,54
79	51,93	47,06	49,72	1,62	0,51	0,43
80	52,52	49,28	52,81	2,36	0,56	0,74
81	45,68	37,87	39,49	1,85	1,93	1,17
82	33,87	26,62	29,97	2,16	0,35	0,46
83	73,40	70,03	71,10	1,32	0,40	0,25
84	78,73	75,25	75,62	0,89	0,35	0,11
85	63,39	62,05	65,39	6,09	0,92	1,69
86	67,14	63,98	65,90	3,06	0,07	0,54
87	66,75	63,72	65,76	3,63	0,56	1,14
88	70,73	67,41	68,82	3,02	0,09	0,46
89	87,78	84,73	84,22	0,20	0,28	0,53
90	73,36	72,95	75,67	7,80	1,58	2,40
Average	57,40	54,80	54,04	1,72	0,89	0,68

Table 1 shows the MMRE value of effort and time development estimation performance of 93 projects in NASA 93. Table 1 shows that the proposed method achieved lower MMRE compared to COCOMO-II and CS. It means that the proposed method found the optimal parameters. So that, the predicted value is closer to the actual effort.

The parameters which used to estimate the value of effort are A and B parameters. The value of A and B will be the balancing parameter for calculating the value of Estimating Effort or Person Month (PM) and exponent E. In addition, the value of PM and E are also influenced by the value of effort multiplier, line of code, and scale factor for each project. By implementing the proposed method in NASA 93, the value of A and B can be optimized well. So that, the proposed method can outperform COCOMO-II and CS model in estimating effort.

The parameters which used to estimate the value of time development are C and D parameters. The value of C and D will be the balancing parameter for calculating the value of estimated time development (TDEV) and exponent F. The influencing parameters of time development estimation are estimated effort, parameter C, and parameter D. As shown in Table 1, the value of proposed method's estimating value is lower than COCOMO and CS in estimating time development. In other words, the most influencing factors are parameter C and D.

Furthermore, not 93 projects are getting the optimal value, there are also some projects which is worse than COCOMO II and CS, such as project 58. By implementing the proposed method in NASA 93, the value of C and D can be failed in project 58. All comparison of time development used a same value of scale factor for each projects. So, the influencing parameters of time development estimation are estimated effort, parameter C, and parameter D. The estimation of effort is vice versa. When there estimated value of proposed method is getting worse than COCOMO II and CS, the balancing value of A and B is being the influencing parameter for that project. In other words, the four coefficients, A, B, C, D being the most influencing parameters in estimating effort and time development. Moreover, each project has its own parameters value such as line of code, effort multiplier, scale factor. So that, the optimized coefficients value can not be good at estimating all projects of NASA 93.

Table 2. Optimized COCOMO-II Coefficient Values.

Coefficient	Value
A	4,631
B	0,81
C	4,1
D	0,3

Table 2 shows the optimized parameter using hybrid cuckoo optimization and harmony search algorithm in NASA 93 dataset. The result of MMRE performance in 93 project is showed in Table 3.

Table 3. Mean of Magnitude Relative Error (MMRE) in 93 projects of NASA 93 dataset

	Effort Estimation	TDEV Estimation
COCOMO II	57,40	1,72
CS	54,8	0,89
CSHS	54,04	0,68

Table 3 shows that the proposed method can obtain lower MMRE in estimating effort and time development compared to COCOMO-II and CS.

6. Conclusion and Future Work

A hybrid model of cuckoo search and harmony search algorithm is proposed for optimizing COCOMO II coefficients. Experimental result was done by using NASA 93 to verify the proposed method's reliability. The experiment showed that the proposed method achieved 54,04% of effort estimation MMRE and 0,68% of time development MMRE which more optimal than COCOMO and CS's estimation. For the future work, this method can be improved to get better result by tuning COCOMO II's cost driver coefficients.

Acknowledgements

Authors give thank to The Indonesia Endowment Fund of Education (LPDP), Ministry of Finance, Indonesia for financial assistance and encouragement.

References

- Rijwani P, Jain S. Enhanced Software Effort Estimation Using Multi Layered Feed Forward Artificial Neural Network Technique. *Procedia Comput Sci.* 2016;89:307-312. doi:10.1016/j.procs.2016.06.073.
- Ghatasheh N, Faris H, Aljarah I, Al-sayyed RMH. Optimizing Software Effort Estimation Models Using Firefly Algorithm. *J Softw Eng Appl.* 2015;8(March):133-142.
- Sarno R, Sidabutar J, Sarwosri. Comparison of different Neural Network architectures for software cost estimation. In: *2015 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*. ; 2015:68-73. doi:10.1109/IC3INA.2015.7377748.
- Zare F, Khademi Zare H, Fallahnezhad MS. Software effort estimation based on the optimal Bayesian belief network. *Appl Soft Comput.* 2016;49:968-980. doi:10.1016/j.asoc.2016.08.004.
- Chalotra S, Sehra SK, Brar YS, Kaur N. Tuning of COCOMO model parameters by using Bee Colony Optimization. *Indian J Sci Technol.* 2015;8(14). doi:10.17485/ijst/2015/v8i14/70010.
- Aljahdali S, Sheta AF. Software effort estimation by tuning COOCMO model parameters using differential evolution. In: *ACS/IEEE International Conference on Computer Systems and Applications - AICCSA 2010*. ; 2010:1-6. doi:10.1109/AICCSA.2010.5586985.
- Sachan RK, Nigam A, Singh A, et al. Optimizing Basic COCOMO Model Using Simplified Genetic Algorithm. *Procedia Comput Sci.* 2016;89:492-498. doi:10.1016/j.procs.2016.06.107.
- Jafari SMS, Ziaaddini F. Optimization of Software Cost Estimation using Harmony Search Algorithm. In: *1st Conference on Swarm Intelligence and Evolutionary Computation (CSIEC)*. ; 2016:131-135. doi:10.1007/978-3-662-47926-1.
- Parwita IMM, Sarno R, Puspaningrum A. Optimization of COCOMO II Coefficients using Cuckoo Optimization Algorithm to Improve The Accuracy of Effort Estimation. In: *International Conference on Information, Communication Technology and System (ICTS), 2016 Int.* ; 2017.
- Sarno R, Sidabutar J, Sarwosri. Improving the Accuracy of COCOMO's Effort Estimation Based on Neural Networks and Fuzzy Logic Model. In: *International Conference on Information, Communication Technology and System.* ; 2015:197-202. doi:10.1109/ICTS.2015.7379898.
- Madheswaran M, Sivakumar D. Enhancement of Prediction Accuracy in Cocomo Model for Software Project Using Neural Network. In: *5th International Conference on Communication and Network Technology.* ; 2014.
- Putri RR, Sarno R, Siahaan D, Ahmadiyah AS, Rochimah S. Accuracy Improvement of the Estimations Effort in Constructive Cost Model II Based on Logic Model of Fuzzy. *Adv Sci Lett.* 2017;23(3):2478-2480. doi:10.1166/asl.2017.8767.
- Gharehchopogh FS, Maleki I, Talebi A. Using hybrid model of Artificial Bee Colony and Genetic Algorithms in Software Cost Estimation. In: *9th International Conference on Application of Information and Communication Technologies (AICT)*. ; 2015:102-106. doi:10.1109/ICAICT.2015.7338526.

14. Miandoab EE, Gharehchopogh FS. A Novel Hybrid Algorithm for Software Cost Estimation Based on Cuckoo Optimization and K-Nearest Neighbors Algorithms. *Eng Technol Appl Sci Res*. 2016;6(3):1018-1022. <http://etasr.com/index.php/ETASR/article/view/701/353>.
15. Maleki I, Ebrahimi L, Gharehchopogh FS. A Hybrid Approach of Firefly and Genetic Algorithms in Software Cost Estimation. In: *MAGNT Research Report*. Vol 2. ; 2014:372-388.
16. Kumar S, Sujan S, Saha K, Conference I, Intelligence C. Advances in Computational Intelligence. In: *International Conference on Computational Intelligence 2015*. Vol 509. ; 2017. doi:10.1007/978-981-10-2525-9.
17. Yang X-S, Deb S. Cuckoo Search via Levy Flights. In: *World Congress on Nature and Biologically Inspired Computing, NABIC*. ; 2009:210-214. doi:10.1109/NABIC.2009.5393690.
18. Geem ZW, Kim JH, Logonathan G V. Harmony search optimization : Application to pipe network design. *Int J Model Simul*. 2002;22(2):125-133. doi:10.1080/02286203.2002.11442233.
19. Sheikholeslami R, Zecchin AC, Zheng F, Talatahari S. A hybrid cuckoo–harmony search algorithm for optimal design of water distribution systems. *J Hydroinformatics*. 2015;(November):544-563. doi:10.2166/hydro.2015.174.
20. Elyasigomari V, Lee DA, Screen HRC, Shaheed MH. Development of a two-stage gene selection method that incorporates a novel hybrid approach using the cuckoo optimization algorithm and harmony search for cancer classification. *J Biomed Inform*. 2017;67:11-20. doi:10.1016/j.jbi.2017.01.016.
21. Feng Y, Wang G, Gao X. A Novel Hybrid Cuckoo Search Algorithm with Global Harmony Search for 0-1 Knapsack Problems. *Int J Comput Intell Syst*. 2016;9(6):1174-1190. doi:10.1080/18756891.2016.1256577.
22. Wang GG, Gandomi AH, Zhao X, Chu HCE. Hybridizing harmony search algorithm with cuckoo search for global numerical optimization. *Soft Comput*. 2016;20(1):273-285. doi:10.1007/s00500-014-1502-7.