Interaction between un updated FR-CG algorithms with optimal Cuckoo algorithm

Rana Z. Al-Kawaz[1], Abbas Y. Al-Bayati[2], Marwan S. Jameel[3]
[1]Department of Mathematics, College of Basic Education, University of Telafer, Iraq
[2]University of Telafer, Iraq
[3]Department of Environmental Technology, College of Environmental Sciences and Technology,
University of Mosul, Iraq

Article Info

ABSTRACT

In this article, we have derived two versions ξ_k and ρ_k were derived from an algorithm based on the first suggested modified Fletcher-Reeves method in the article for the two-term CG method and another term to get a downward search towards the function minimum point with the search for an inaccurate line and we have proved rapprochement. These two algorithms combined with the Cuckoo algorithm to achieve a remarkable performance in reducing the number of repetitions in order to reach the minimization of 10 functions is unconstrained in the numerical results.

Corresponding Author:

Rana Z. Al-Kawaz,
Department of Mathematics,
College of Basic Education,
University of Telafer, Mosul, Iraq.
Email: rana.alkawaz@yahoo.com

1. INTRODUCTION

Previously, the issue of large-dimensional optimization was solved using widely derived methods such as CG techniques, especially when general functions existed. This method is designed for unconstrained optimization in the following manner:

$$min\, f\,(x) \text{ where Domain} = R^n \text{ and Range} = R \tag{1}$$

The variable here is a vector, and function f is a continuous and nonlinear function. The approach to the associated gradient methods is a repetitive method that defines the function at the iterative points i.e. f at x_k $(f_k = f(x_k))$, The approach to the associated gradient methods is a repetitive method that defines the function at the iterative points $g_k = g(x_k)$ and the second derivative of the function is the formula $G_k = G(x_k)$ in the iteration k (Hessian matrix). This method generates a sequential repeat of the step calculated from the following formula:

$$x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

such that $k \geq 0$ and α_k is the measure of the step between each successive point [1]. The strong Wolfe (SWC) line search is one of the important methods to calculate the step scale when solving general functions and can be defined α_k as a form [2]:

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \\ |d_k^T g_{k+1}| \leq -\sigma d_k^T g_k \ 0 \leq \delta \leq \sigma \leq 1 \end{cases} \tag{3}$$

Multiplied by the search direction in which it is:

$$d_{k+1} = \begin{cases} -g_{k+1} \ \text{if } k = 0, \\ -g_{k+1} + \beta_k d_k, \text{if } k > 0, \end{cases} \tag{4}$$

The CG method depends mainly on the selection of a parameter between 0 and 1 which is the parameter β_k to illustrate its importance and types, its main definitions can be read [3]."The well-known formulas for β_k are the Fletcher-Reeves (FR) [4], Hestenes-Stiefel (HS) [5] and Polak-Ribière (PR) [6, 7] formulas, which are specified by:

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \text{ , (Fletcher Reeves (FR),1964)}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \text{ , (Hestenes Stiefel (HS),1952)}$$

Such that $\|.\|$ known as the two-point Euclidean distance. The measure of the extent of accompaniment is through the achievement of one of the Conjugacy conditions, which ranges from the usual to the general and the general ones in the following formulas:

$$d_{k+1}^T y_k = 0. \tag{5}$$

$$d_{k+1}^T y_k = -g_{k+1}^T s_k \tag{6}$$

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k. \tag{7}$$

S.t. t is a positive scalar, the last version is the widest of the three formulas, from which the rest can be derived and called a formula Dai and Liao [8]. To facilitate understanding of these formulas $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$. Many kind of research have discussed the convergence of methods used for parameters FR, HS and PR s.t. Zoutendijk [2] and Al-Baali [9]. Based on these theories in the previous article, many researchers have developed and updated the formulas of the correlation parameter as in [10-18].

Swarm intelligence algorithms are of great importance at this time, namely, Metaheuristics algorithms, for example the Cuckoo bird algorithm, where cuckoo birds behave poorly to use other birds nests and lay their eggs. [19] Birds that host cuckoo eggs do two things, either to leave their eggs or to leave those nests. While the cuckoo adjusts its eggs by simulating the color and pattern of host eggs. There are rules that this Cuckoo algorithm follows in its behavior ((1) Distribute eggs randomly on other birds nests each time. (2) Breed future generations at the best nest (good eggs). (3) Make sure the eggs that the cuckoo had in the host birds fixed the number for each nest with the probability pα ∈ [0, 1], i.e. this method, give new random solutions and replace the nest host.

According to the rules (CS) there can be one solution and there can be a set of solutions according to the number of eggs in the nest. According to this conclusion, each bird is likely to place only one egg in the nest in the same original form, giving us that each nest contains eggs of multiple traits representing a set of solutions, as a general idea. Mathematically, these types of issues are reduced to the maximum transformation of problems with respect to the exact equation. Here the objective function referred to as the fitness function [20]. For a random search, the animal search path is randomly chosen without any discrimination as the next step depends on (location - current status - likely to move to your next location). The implicit trend depends on the probability of the mathematical model. The cuckoo algorithm can be written clearly. It is the search for x that reduces the function of the target function f (x) and through these rules, we conclude that for cuckoo nest i, the solution for the new generation can be defined as [21]:

$$x_{k+1}^i = x_k^i + \tau_1 \oplus \text{Lévy} (\tau_2) \tag{8}$$

Let it be $(\tau_1 = 1.5)$ the amount of possible flight Lévy is possible, $\oplus$ multiplication symbol meaning double entries and After the search went beyond random search behavior in the nature of the animals when the characteristics of the behavior converge Lévy flight. We write the distribution:

$$\text{Lévy} \sim N = k^{-\tau_2}, 0 < \tau_2 \leq 3 \tag{9}$$

N denotes that it follows the natural distribution by the number of frequencies k is. How many times do we produce the numbers randomly, obey Lévy flights, Calculate the amount and direction of the step in a random way using the Levy distribution contained in the two search papers [22-23]. So, it is clear to us that if the Cuckoo's egg is very similar to host eggs, it may lead to the disappearance of this Cuckoo's egg, so the fitness must be linked to different solutions and it is recommended that a random path be conducted in a biased manner with some random steps size [24-25].

We will organize our search as follows: In Section 2, in fact, will derive two new formulas for CG-algorithm with a descent characteristic of these two algorithms. In Section 3, the theoretical side and derive the attributes of the global convergence of the two new CG-algorithm under some assumptions. In Section 4, combine the CG-algorithm with the Cuckoo's-algorithm to reinforce each other. In Section 5, to assign the search to the numerical side by applying the two integrated algorithms on 10 unconstrained functions and comparing them with the Cuckoo algorithm.

2.   TWO NEW FORMULAS FOR CG

In this part of the article, we will give an update of the parameter of β_k^{FR} in two forms once in two and three times with a given derivation of the two formulas. All the proposed parameters were derived using the normal conjugacy condition (5) as in the following equations

$$d_{k+1}^{New} = -\frac{1}{\rho_k^2 - 1} g_{k+1} + \beta_k^{FR} s_k \tag{10}$$

By dealing with the normal conjugacy condition (5), and when multiplying the two ends of (12) by y_k:

$$y_k^T d_{k+1} = -\frac{1}{\rho_k^2 - 1} y_k^T g_{k+1} + \beta_k^{FR} y_k^T s_k$$

$$0 = -\frac{1}{\rho_k^2 - 1} y_k^T g_{k+1} + \beta_k^{FR} y_k^T s_k$$

$$\frac{1}{\rho_k^2 - 1} y_k^T g_{k+1} = \beta_k^{FR} y_k^T s_k$$

$$\frac{y_k^T g_{k+1}}{\beta_k^{FR} y_k^T s_k} = \rho_k^2 - 1$$

$$\rho_k^2 = \frac{\beta_k^{HS}}{\beta_k^{FR}} + 1$$

And when taking the square root of the parties hence:

$$\rho_k^{new} = \mp\sqrt{\frac{\beta_k^{HS} + \beta_k^{FR}}{\beta_k^{FR}}} \tag{11}$$

Now we can consider that the formulas (10) and (11) are the first updated algorithm, and to develop them we add a third term to the new search direction of the format:

$$d_{k+1}^{New} = -\frac{1}{\rho_k^2 - 1} g_{k+1} + \beta_k^{FR} s_k - \xi_k^T y_k \tag{12}$$

In the same way as the derivation of the parameter ρ_k we derive the third term parameter as well:

$$\xi_k^{new} = \frac{\beta_k^{FR}}{\|y_k\|^2} [(\beta_k^{HS} + \beta_k^{FR}) y_k^T s_k + y_k^T g_{k+1}] \tag{13}$$

The last two in (12-13) represent Formula 2 for modernization the formula of β_k^{FR}.

3. PROVE CONVERGENCE OF THE CG ALGORITHM

To prove in this part of the article that the new algorithms in the use of mathematical tools in theoretical analysis led to a comprehensive convergence and gave efficiency when compared to the basic algorithm under the conditions of the line of search and retrieval, now we provide the basics of convergence analysis used, For any theory prepared, suppose that, g_k never equal to zero for all $k \geq 1$, otherwise, a stationary point was found, including the following basics on the approved function:

3.1. Assumption
When x_1 is exists, then the level set $S = \{x: f(x) \leq f(x_1)\}$ is bounded, namely there exists a number;

$$(B > 0) \text{ such that} \|x\| \leq B, \forall x \in S \tag{14}$$

In some N neighborhoods of S, f it can be distinguished continuously differentiable and g is Lipschitz continuously added, namely, there exists a constant $L \geq 0$ such that

$$\|g(x) - g(x_k)\| \leq L\|x - x_k\|, \forall x, x_k \in N \tag{15}$$

3.2. Theorem (descent condition-2)
Let x_{k+1} and d_{k+1} be generated by two new Algorithm (10-13), respectively, and let α_k be obtained by the SWC (3), then the direction holds s. t.

$$g_{k+1}^T d_{k+} \leq -c\|g_{k+1}\|^2, \forall k \geq 1 \tag{16}$$

Proof:
Multiplying (12) by g_{K+1}, we have:

$$d_{k+1}^T g_{k+1} = -\frac{1}{\rho_k^2 - 1}\|g_{k+1}\|^2 + \beta_k^{FR} s_k^T g_{k+1} - \xi_k\, y_k^T g_{k+1}$$

$$d_{k+1}^T g_{k+1} = -\left[\frac{1}{\rho_k^2 - 1} + \frac{s_k^T g_{k+1}}{\|g_k\|^2}\right]\|g_{k+1}\|^2 - \xi_k\, y_k^T g_{k+1}$$

Where $g_{k+1}^T s_k = y_k^T s_k + g_k^T s_k > g_k^T s_k$. For the first algorithm, when eliminating the last term and applying the last condition to it, we get sufficient descent for this algorithm as follows:

$$d_{k+1}^T g_{k+1} \leq -c_1\|g_{k+1}\|^2$$

Where $c = \left[\frac{1}{\rho_k^2 - 1} + 1\right] > 0, and\ \frac{1}{\rho_k^2 - 1} > 0$

Let's complete the descent proof of the second algorithm, in addition, and using Powell restart criteria

$$y_k^T g_{k+1} = g_{k+1}^T g_k - g_k^T g_{k+1} \geq (1 - \psi)\|g_{k+1}\|^2$$

Then

$$d_{k+1}^T g_{k+1} \leq -\left[\frac{1}{\rho_k^2 - 1} + 1\right]\|g_{k+1}\|^2 - \frac{(1-\psi)^2\|g_{k+1}\|^6}{\|y_k\|^2\|g_k\|^2}\left[2 + \frac{y_k^T s_k}{\|g_k\|^2}\right]$$

$$d_{k+1}^T g_{k+1} \leq -\left[\left[\frac{1}{\rho_k^2 - 1} + 1\right] - \frac{(1-\psi)^2\|g_{k+1}\|^4}{\|y_k\|^2\|g_k\|^2}\left[2 + \frac{y_k^T s_k}{\|g_k\|^2}\right]\right]\|g_{k+1}\|^2$$

Let

$$c_2 = \left[\left[\frac{1}{\rho_k^2 - 1} + 1\right] - \frac{(1-\psi)^2\|g_{k+1}\|^4}{\|y_k\|^2\|g_k\|^2}\left[2 + \frac{y_k^T s_k}{\|g_k\|^2}\right]\right] > 0$$

$\frac{1}{\rho_k^2 - 1} > 0$ and $y_k^T s_k > 0$

Hence,

$$d_{k+1}^T g_{k+1} \leq -c_2 \|g_{k+1}\|^2.$$

3.3. Theorem

Let Assumptions (3.1) A holds and consider any CG-algorithm (2-4), where d_{k+1} is a satisfied (20) and α_k is obtained by (3), if

$$\sum_{k \geq 1} \frac{1}{\|d_k\|^2} < +\infty \tag{17}$$

Then, we have

$$\lim_{k \to \infty} \inf \|g_k\| = 0 \tag{18}$$

Let us now give the theory of global convergence using conditions in previous theories of proof:

3.4. Theorem

Let Assumptions (3.1) B hold. Assume that $0 < \rho_k , \xi_k < 1$, for every, k≥0, there exists a positive constant c_i, $(i = 1,2)$ as the constants γ_1 and γ_2 such that $\gamma_1 \leq \|g_k\| \leq \gamma_2$. Then, the two new scheme and α_k is determined by the SWC search, either $g_k = 0$ for some k or $\lim_{k \to \infty} \inf \|g_k\| = 0$

Proof:

Because the descent condition holds, we have $\|d_k\| \neq 0$.using the Lipchitz condition

$\|y_k\| = \|g_{k+1} - g_k\| \leq L\|s_k\|$ then,

$$|\rho_k^2| = \left| \frac{\beta_k^{HS}}{\beta_k^{FR}} + 1 \right| \leq \frac{\|y_k\|\|g_{k+1}\|}{\|s_k\|\|y_k\|} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} + 1 \leq \frac{1}{\|s_k\|} \frac{\|g_k\|^2}{\|g_{k+1}\|} + 1$$

$$\leq \frac{\gamma_1^2}{\gamma_2} \frac{1}{D} + 1$$

D as knew in assumption and $|\beta_k^{FR}| \leq \frac{\gamma_2^2}{\gamma_1^2} \equiv E$,

$$|\xi_k^{new}| \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2 \|y_k\|^2} \left[\left(\frac{\|y_k\|\|g_{k+1}\|}{\|s_k\|\|y_k\|} + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \right) \|y_k\|\|s_k\| + \|y_k\|\|g_{k+1}\| \right]$$

$$|\xi_k^{new}| \leq \frac{\gamma_2^2}{\gamma_1^2 L^2} \left[\left(\frac{\gamma_2}{\|s_k\|} + \frac{\gamma_2^2}{\gamma_1^2} \right) L\|s_k\| + L\gamma_2 \right]$$

$$|\xi_k^{new}| \leq \frac{\gamma_2^2}{\gamma_1^2 L} \left[\left(\frac{\gamma_2}{D} + \frac{\gamma_2^2}{\gamma_1^2} \right) D + \gamma_2 \right] \equiv F$$

the direction

$$\|d_{k+1}^{New}\| \leq \left| \frac{D\gamma_2}{\gamma_1^2} \right| \gamma_2 + ED + FL$$

$$\leq \frac{\gamma_2^2 D}{\gamma_1^2} + ED + FL$$

This implies $0 < \sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \; \sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \frac{1}{c^2} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$

3.5. New 1 (2-Term) CG-Algorithm:

Step 1: Inserts: initialized Cuckoo parameter (pα= 0.25, N-IterTotal=1000)
Give initial CG parameter: variable $x_0 \in R^n$, $\delta \in [0,0.5]$ and $\sigma \in [\delta, 1]$. Let k = 0, $d_0 = -g_0$.
Step 2: The best amount of step to algorithm results: set α_k from SWC in (3), calculate the parameters used by the new 1 search direction (10-11).
Step 3: Find the new point happened x_{k+1} as (2-4).
Step 4: Using this point in (9) to evaluate the new Lévy flight.
Step 5: Evaluate the new solution (8).
Step 6: If the value of the total number of duplicates has ended, stop and print the new value is the best and not adjusted in
k = k+1 and go to Step (2).

3.6. New 2 (3-Term) CG-Algorithm:

Step 1: Inserts: initialized Cuckoo parameter (pα= 0.25, N-IterTotal=1000)
Give initial CG parameter: variable $x_0 \in R^n$, $\delta \in [0,0.5]$ and $\sigma \in [\delta, 1]$. Let k = 0, $d_0 = -g_0$.
Step 2: The best amount of step to algorithm results: set α_k from SWC in (3), calculate the parameters used by the new 1 search direction (12-13).
Step 3: Find the new point happened x_{k+1} as (2-4).
Step 4: Using this point in (9) to evaluate the new Lévy flight.
Step 5: Evaluate the new solution (8).
Step 6: If the value of the total number of duplicates has ended, stop and print the new value is the best and not adjusted in
k = k+1 and go to Step (2).

4. NUMERICAL RESULTS:

       In this part of the article, the new algorithms were implemented on ten test functions taken from the source Jamil and Yang [26] and we used a program provided by Yang [27]. It has been modified to fit with the new algorithms and implemented the program within the MATLAB R2018b during PC CORE I5 laptop. The results are categorized into 3 main categories. The results of the new algorithms were compared with the original Cuckoo algorithm (OCA) as follows:
1- n=25 &NOI=5000
2- n=50 &NOI=100000
3- n=250 &NOI=500000

And finding the f_{min} of the function that showed the following Table 1 the efficiency of the new algorithms.
       If we observe the tables closely and we understand the values, we note the superiority of the performance of the new algorithms on the algorithm of the patch to find the miniaturization of the test functions used within the article in the following numbers:
a)   The first algorithm (two-term) has exceeded 10 times the cuckoo algorithm and different frequencies.
b)   The second algorithm (three-term) has exceeded 20 times the cuckoo algorithm and with a different frequency.

Table 1. The performance of the two new CCG algorithms compared to the original algorithm

Function name	Dimension	f_{min} OCA Vs. (CCG-1 - CCG-2)
Sphere	n=25 & NOI=5000	8.2807e-08 Vs. (5.316312e-19 - 1.704955e-18)
	n=50 & NOI=100000	1.0777e-06 Vs. (9.330753e-21 - 1.132487e-19)
	n=250 & NOI=500000	9.8068e-05 Vs. (3.237586e-19 - 2.036010e-19)
ALPINE 1	n=25 & NOI=5000	0.4933 Vs. (1.389594e-16 - 2.460353e-19)
	n=50 & NOI=100000	0.7902 Vs. (2.354100e-18 - 3.567730e-20)
	n=250 & NOI=500000	0.8511 Vs. (2.208516e-21 - 1.045163e-18)
BROWN	n=25 & NOI=5000	3.2250e-04 Vs. (1.527760e-18 - 1.839775e-18)
	n=50 & NOI=100000	0.1256 Vs. (1.037208e-18 - 7.686261e-19)
	n=250 & NOI=500000	0.939 Vs. (1.585121e-21 - 1.192826e-19)
EXPONENTIAL	n=25 & NOI=5000	-1.0000 Vs. (2.019619e-19 - 8.954126e-18)
	n=50 & NOI=100000	-1.0000 Vs. (2.614323e-19 - 1.852968e-19)
	n=250 & NOI=500000	-0.9999 Vs. (7.179282e-20 - 5.300398e-20)

Function name	Dimension	f_{min} OCA Vs. (CCG-1 - CCG-2)
GRIEWANK	n=25 & NOI=5000	3.4194e-07 Vs. (5.667865e-19 - 2.620217e-18)
	n=50 & NOI=100000	1.9925e-04 Vs. (1.730804e-19 - 2.656963e-18)
	n=250 & NOI=500000	9.5789e-04 Vs. (1.665153e-18 - 2.649322e-19)
PERIODIC	n=25 & NOI=5000	1.0196 Vs. (1.560061e-17 - 7.818540e-21)
	n=50 & NOI=100000	1.0150 Vs. (2.073343e-18 - 1.463280e-19)
	n=250 & NOI=500000	1.0145 Vs. (1.168592e-19 - 9.218664e-20)
POWELL SUM	n=25 & NOI=5000	2.5433e-08 Vs. (1.672876e-18 - 1.553991e-19)
	n=50 & NOI=100000	4.8484e-06 Vs. (4.516643e-19 - 2.918082e-19)
	n=250 & NOI=500000	0.0139 Vs. (8.660050e-19 - 8.920962e-20)
SALOMON	n=25 & NOI=5000	0.2999 Vs. (5.442767e-18 - 2.685704e-18)
	n=50 & NOI=100000	0.3999 Vs. (3.940597e-19 - 3.792677e-19)
	n=250 & NOI=500000	0.4999 Vs. (4.527411e-20 - 7.371142e-20)
SCHWEFEL 2.23	n=25 & NOI=5000	4.0638e-12 Vs. (1.844583e-17 - 3.526535e-18)
	n=50 & NOI=100000	1.0075e-18 Vs. (2.361283e-21 - 6.668129e-22)
	n=250 & NOI=500000	1.1731e-10 Vs. (7.059524e-24 - 1.478208e-20)
ROSENBROCK	n=25 & NOI=5000	19.781 Vs. (2.511745e-17 - 7.962472e-18)
	n=50 & NOI=100000	20.9467 Vs. (6.798510e-19 - 6.475936e-17)
	n=250 & NOI=500000	23.6746 Vs. (1.074577e-19 - 8.996584e-24)

5. CONCLUSIONS

We note that from our numerical results that the two new algorithms perform better than the basic algorithm (Cuckoo) by 90%, which indicates that it is not possible to use these two algorithms mixed with each other to improve the performance of the associated gradient algorithm through the cuckoo algorithm using some distinct techniques within the two methods.

REFERENCES

[1] L. Zhang, W. J. Zhou and D. H. Li, "Global Convergence of a Modified Fletcher-Reeves Conjugate Method with Armijo-Type Line Search," *Numerische Mathematik*, vol. 104, no. 4, pp. 561-572, 2006.
[2] G. Zoutendijk, "Nonlinear Programming, Computational Methods", in *Integer and Nonlinear Programming, Journal of Abadie (Ed.)*, North-Holland: Amsterdam, pp. 37–86, 1970.
[3] N Andrei, "Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization" *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 34, no. 2, 2011.
[4] R. Fletcher, C. Reeves, "Function minimization by conjugate gradients," *Comput. Journal*, vol. 7, pp. 149-154, 1964.
[5] M. R. Hestenes, E. L. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems", *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409-432, 1952.
[6] B. Polak and G. Ribiere, "Note Surla Convergence des Méthodes de Directions Conjuguées", *Revue Francaise d'Informatique et de Recherche Opérationnelle*, vol. 16, no. 1, pp. 35-43, 1969.
[7] B. T. Polyak, "The Conjugate Gradient Method in Extreme Problems", *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94-112, 1969.
[8] Y. H. Dai and L. Z. Liao, "New conjugacy conditions and related nonlinear conjugate gradient methods", *Applied Mathematics and Optimization*, vol. 43, no. 1, pp. 87-101, 2001.
[9] M. Al-Baali, "Descent Property and Global Convergence of the Fletcher-Reeves Method with Inexact Line Search", *IMA Journal of Numerical Analysis*, vol. 5, pp. 121– 124, 1985.
[10] H. Y., Najm, E. T. Hamed, and H. I. Ahmed. "Global convergence of conjugate gradient method in unconstrained optimization problems," *AIP Conference Proceedings*. vol. 2086. no. 1. AIP Publishing LLC, 2019.
[11] A. Zhou, Z. Zhu, H. Fan, Q. Qing, "Three New Hybrid Conjugate Gradient Methods for Optimization", *Applied Mathematics*, vol. 2, pp. 303-308, 2011.
[12] N. S. Mohamed, M. Mamat, M. Rivaie, and S. M. Shaharudin, "A new hybrid coefficient of conjugate gradient method," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, (3), pp. 1454-1463, 2020.
[13] M. K. Dauda, M. Mamat, M. A. Mohamed and N. S. A. Hamzah, "Hybrid conjugate gradient parameter for solving symmetric systems of nonlinear equations," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 16, pp. 539-543, 2019.
[14] J. Z. Zhang, C. X. Xu, "Properties and Numerical Performance of Quasi-Newton Methods with Modified Quasi-Newton Equation", Journal of Comput. Appl. Math., vol. 137, pp. 269-278, 2001.
[15] D. H. Li, M. Fukushima, "A Modified BFGS Method and Its Global Convergence Non-Convex Minimization", Journal of Comput. Appl. Math., vol. 129, pp. 15-35, 2001.
[16] H. I. Ahmed, R. Z. Al-Kawaz and A. Y. Al-Bayati, "Spectral Three-Term Constrained Conjugate Gradient Algorithm for Function Minimizations," *Journal of Applied Mathematics*, vol. 2019, 2019.
[17] A. Y. Al-Bayati, M. S. Al-Jameel, "New Scaled Proposed Formulas for Conjugate Gradient Methods in Unconstrained Optimization", AL-Rafidain J. of Computer Science and Mathematics, Mosul, Iraq, vol. 11, no. 2, pp.25-46, 2014.

[18] B. A. Hassan, "A new formula for conjugate parameter computation based on the quadratic model," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, (3), pp. 954-961, 2019.

[19] R. B. Payne and M. D. Sorensen, "The Cuckoos," *Oxford*, UK: Oxford University Press. 2005.

[20] I. Jr. Fister, D. Fister, I. Fister, "A comprehensive review of cuckoo search: variants and hybrids", International Journal of Mathematical Modeling and Numerical Optimization, vol. 4, no. 4, pp.387–409, 2013.

[21] A. H. Gandomi, X. S. Yang and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems", *Eng. Comput*., vol. 29, no. 1, pp. 17–35, 2013. DOI: 10.1007/s00366-011-0241-y.

[22] H. L. Chen, B. Yu, H. L. Zhou, Z. Meng, "Improved Cuckoo Search Algorithm for Solving Inverse Geometry Heat Conduction Problems", Heat Transfer Engineering, Taylor and Francis, vol. 0, no. 0, pp.1–13, 2018.

[23] I. Pavlyukevich, "Lévy flights, non-local search and simulated annealing", J. Comput. Phys., vol. 226, pp. 1830–1844, 2007. DOI: 10.1016/j.jcp.2007.06.008.

[24] X. S. Yang, S. Deb, "Cuckoo search via Levy flights", in: Proc. of World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), India, Publications, USA, pp. 210-214, 2009.

[25] R. Z. Al-Kawaz, H. I. Ahmed, A. Y. Al-Bayati, "An Efficient Modified Cuckoo Search Optimization with the Rational Barrier Function," *International Journal of Mathematics and Computer Applications Research (IJMCAR)*, TJPRC. vol. 8, Issue 5, pp. 13-24, 2018.

[26] M. Jamil, X. S. Yang, "A Literature Survey of Benchmark Functions for Global Optimization Problems", Int. J. Math. Model. Numer. Optim, vol. 4, pp. 150–194, 2013.

[27] X. S. Yang, S. Deb, "Engineering optimization by cuckoo search", Int. J. Mathematical Modelling and Numerical Optimization, vol. 1, no. 4, pp. 330-343, 2010. http://arxiv.org/PS_cache/arxiv/pdf/1005/1005.2908v2.pdf [23]