Using Modified Conjugate Gradient Method to Improve SCA

Ayad Hamad Khalaf¹, Prof. Dr. Ban Ahmed Mitras²

¹ Student, department of mathematics, college of computer sciences & mathematics, Mosul University. Email: ayadlohebe@Gmail.com
² Prof. Dr. department of mathematics, college of computer sciences & mathematics, Mosul University. Email: dr.banah.mitras@Gmail.com

Abstract: This research is to improve Sine - Cosine Algorithm (SCA) that is like any other intelligent techniques that encounter some problem such as slow convergence and the dropping in local solution. To overcome these problems, SCA has been developed and improved through three directions, First: Hybrid of SCA with Modified conjugate gradient method (MCG) that has improved through that derivation of parameter of new conjugate factor ($\beta^{new}$) and attest its characteristic such as descent and global to construct improve algorithm called SCA-MCG. The second direction was a hybrid of SCA with classic optimization methods such as conjugate gradient (CG) algorithm to construct SCA-CG, and parallel Tangent (PT) algorithm to construct SCA-PT. Third combining both previous methods, using the Hybrid value with SCA to construct SCA-CG-PT Algorithm of high quality accounts in all directions mentioned above. To improve the initial population which randomly generated by using excellent characteristics of MCG-CG-PT as well as using this improvement as initial population for SCA. Numerical results have proved the efficiency of improved Algorithm and the results was excellent if we compared with SCA. In addition, we got optimum global values for most functions by achieving functions minimum.

Keyword: SCA algorithm, meta-heuristic algorithms, conjugate gradient and PARTAN methods

1. Introduction:

Optimization refers to the finding optimum values of the given system facts, for all possible values. In mathematics, it means to find minimum or maximum value of a function contains a certain number of variants. It can found in all fields of study that seek to develop basic optimization techniques so it is of a high importance for most researchers in their works. Optimization started in 1960 through many directions and methods through 2 main parts of algorithms, the first is Deterministic, and the other is Stochastic. Most of classical algorithms are deterministic, such as CG, PARTAN, QN, and others. Most of them based on slope or what called derivation, (derivation base algorithm). The second part of algorithm, Stochastic that is divided into Heuristic and meta-heuristic. It is important to mention that, recent trends of study refer to the lack of certain definition of these Heuristic and Meta Heuristic. (Glover 1986).
SCA is a Heuristic and inspired by sine and cosine functions. It suggested by (Seyedali Marjali 2016) to solve optimization and apply it to improve airplanes' performance [9]. Many improvement and modifications as well as Hybrid suggested. In 2017, (simye ,Busra,Pakite) present a study about Constructed optimization problems using SCA[6] . Same year witnessed presenting another study of a Hybrid of SCA to solve global optimization problems by (R.M,rizk Allah) [13]. In 2018 Zhiliujun , Chiver, et al) present a study about Modifuing SCA based on search of circular uninvited adjunct. In the same year, ( Ramzy Ali , Dunisis) present a study about Chaotic SCA[12] . Finally, in 2019 (yasmin, R.sindhu et. al.) present a study of SCA Hybrid using biogeography for problems of choosing merit [14]. In the same year (mohoub, Mohamed, et al.) present a study about improving SCA to choose merit in sorting texts [2]. Two researchers (Lalit, Kusum) present a research paper about choosing merit [8]. (Chandrasekaran), also, present a study to improve SCA on the problem of sending Dynamic economy [3]. (Gholizadeh1, S. & Sojoudizadeh) present a study to modified sine cosine algorithm for Sizing Optimization of Truss Structures [7]. With an explanation of CG method to show its characteristic. Since it is one of the classical methods and use it in generating initial community used with SCA, using its characteristics to get optimum and global solution. PT and its uses with SCA had referred to also. After checking results, we made third modification by combining CG and PT with SCA. The numerical results were better when applied on special functions. Finally, new conjugate factor had derived, and then its globalization and slope was a tested.

It show efficiency when used in Hybrid SCA plus combining suggested and Classical and Heuristic to produce improved and Hybrid algorithm of high Characteristic tested on a set of special functions. The problem of the research focused on finding global optimum solutions for optimization problems to get rid of slow convergence, and fall in local solutions.

The study aims at presenting improved algorithm that hybrid of sine-cosine algorithm SCA with a set of classical algorithms named as SCA-MCG, SCA-CG, SCA-PT and SCA-CG-PT.

2. Conjugate Gradient Method:

In unconstrained optimization, we minimize an objective function that depends on real variables with no restrictions on the values of these variables. The unconstrained optimization problem is:

\[ \text{Min} \quad f(x) : x \in \mathbb{R}^n, \]

(1)

Where \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is a continuously differentiable function, bounded from down. A nonlinear conjugate gradient method generates a sequence \( \{x_k\}, k \) is integer number, \( k \geq 0 \). Starting from an initial point \( x_0 \), the value of \( x_k \) calculate by the following equation:

\[ x_{k+1} = x_k + \lambda_k d_k, \]

(2)

Where the positive step size \( \lambda_k > 0 \) obtained by a line search and the directions \( d_k \) generated as:

\[ d_{k+1} = -g_{k+1} + \beta_k d_k, \]

(3)
Where \( d_0 = -g_0 \), the value of \( \beta_k \) is determine according to the algorithm of Conjugate Gradient (CG), and its known as a conjugate gradient parameter, \( s_k = x_{k+1} - x_k \) and \( g_k = \nabla f(x_k) = f'(x_k) \), consider \( ||\| \) is the Euclidean norm, and \( y_k = g_{k+1} - g_k \). The termination conditions for the conjugate gradient line search often based on some version of the Wolfe conditions. The standard Wolfe conditions [4]:

\[
\begin{align*}
 f(x_k + \lambda_k d_k) - f(x_k) &\leq \rho \lambda_k g_k^T d_k, \\
g(x_k + \lambda_k d_k)^T d_k &\geq \sigma g_k^T d_k.
\end{align*}
\]

Where \( d_k \) is a descent search direction and \( 0 < \rho \leq \sigma < 1 \), where \( \beta_k \) defined by one of the following formulas:

\[
\begin{align*}
\beta_k^{(HS)} &= \frac{y_k^T g_{k+1}}{y_k^T d_k}; & \beta_k^{(FR)} &= \frac{g_{k+1}^T g_k g_k^T g_{k+1}}{g_k^T g_k}; & \beta_k^{(PRP)} &= \frac{y_k^T g_{k+1}}{g_k^T g_k} \\
\beta_k^{(CD)} &= -\frac{g_{k+1}^T g_k g_{k+1}}{y_k^T d_k}; & \beta_k^{(LS)} &= -\frac{y_k^T g_{k+1}}{g_k^T d_k}; & \beta_k^{(DY)} &= \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k}.
\end{align*}
\]

Al-Bayati and Al-Assady In (Al-Bayati and Al-Assady, 1986) proposed three forms for the scalar \( \beta_k \) defined by [1]:

\[
\begin{align*}
\beta_k^{AB1} &= \frac{\|y_k\|^2}{\|g_k\|^2}; & \beta_k^{AB2} &= -\frac{\|y_k\|^2}{d_k^T g_k}; & \beta_k^{AB3} &= \frac{\|y_k\|^2}{d_k^T y_k}.
\end{align*}
\]

3. Extension Dai and Yuan Method:

Yabe and Sakaiwa in 2005 extended the Dai and Yuan method as [4]:

\[
\beta_k = \frac{\|g_{k+1}\|^2}{\tau_{k+1}}
\]

Where \( \tau_{k+1} \) be a positive parameter.

By setting \( \tau_k = d_k^T y_k \) formula (9) reduce to this DY method as:

\[
\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}
\]
4. Proposed A New Conjugacy Coefficient:

We have the quasi-Newton condition

\[ y_k = G_k s_k \]  
(11)

Where \( G_k = \frac{\partial^2 f}{\partial x_k^2} \) is the Hessian Matrix

We multiply both sides of equation (11) by \( s_k \) and we get

\[ [y_k = G_k s_k] \cdot s_k \]  
(12)

\[ \Rightarrow y_k^T s_k = G s_k^T s_k \]  
(13)

\[ G = \frac{y_k^T s_k}{\|s_k\|^2} I_{n,n} \]  
(14)

Where I is the identity matrix

Let \( d_{k+1}^N = -G_k^{-1} g_{k+1} \)  
(15)

Eq. (15) is the Newton direction. From eq.(15) and (15) we get:

\[ d_{k+1}^N = -\frac{y_k^T s_k}{\|s_k\|^2} g_{k+1} \]  
(16)

Multiply both sides of equation (16) by \( y_k^T \) and we get

\[ y_k^T d_{k+1}^N = \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1} \]  
(17)

\[ \Rightarrow y_k^T d_{k+1}^{CG} = -y_k^T g_{k+1} + \beta_k d_k^T y_k \]  
(18)

From (17) and (18) we have:

\[ -y_k^T g_{k+1} + \beta_k d_k^T y_k = \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1} \]  
(19)
We assume that $\beta_k = \beta_k^{(DY)} = \frac{g_{k+1}^T g_{k+1}}{y_k^T d_k}$

Then we have

$$- y_k^T g_{k+1} + \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T y_k = - \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1}$$

From eq. (9) we get:

$$- y_k^T g_{k+1} + \beta_k \tau_k = - \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1}$$

Then, we have

$$\beta_k = \frac{- \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1} + y_k^T g_{k+1}}{\tau_k}$$

Since $\tau_{k+1} > 0$ then from [5], we have: $\tau_k = \lambda = \left[ \frac{\|s_k\|^2}{2(f_k - f_{k+1}) + 2g_{k+1}^T s_k} \right]$ then:

$$\beta_k = \left( - \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1} + y_k^T g_{k+1} \right) + \frac{\|s_k\|^2}{2(f_k - f_{k+1}) + 2g_{k+1}^T s_k}$$

$$\beta_k = \left( - \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] y_k^T g_{k+1} + y_k^T g_{k+1} \right) \frac{2(f_k - f_{k+1}) + 2g_{k+1}^T s_k}{\|s_k\|^2}$$

$$\beta_k = \left( 1 - \left[ \frac{y_k^T s_k}{\|s_k\|^2} \right] \right) y_k^T g_{k+1} \frac{2(f_k - f_{k+1}) + 2g_{k+1}^T s_k}{\|s_k\|^2}$$

Let $A = f_k - f_{k+1}$ then:
\[ \beta_k = \begin{pmatrix} 1 - \frac{y_k^T s_k}{\|s_k\|^2} \end{pmatrix} \cdot \frac{y_k^T g_{k+1} \left(2A + 2g_{k+1}^T s_k\right)}{\|s_k\|^2} \]  \hspace{1cm} (26)

Or

\[ \beta_k = \frac{1}{\|s_k\|^2} \left(1 - \frac{y_k^T s_k}{\|s_k\|^2}\right) \cdot \frac{y_k^T g_{k+1} \left(2A + 2g_{k+1}^T s_k\right)}{\|s_k\|^2} \]  \hspace{1cm} (27)

4.1 Outlines of the Proposed Algorithm:

Step (1): The initial step: We select starting point \( x_0 \in \mathbb{R}^n \), and we select the accuracy solution \( \varepsilon > 0 \) is a small positive real number and we find \( d_0 = -g_k \), \( \lambda_0 = \text{Minary}(g_0) \), and we set \( k = 0 \).

Step (2): The convergence test: If \( \|g_k\| \leq \varepsilon \) then stop and set the optimal solution is \( x_k \). Else, go to step (3).

Step (3): The line search: We compute the value of \( \lambda_k \) by Cubic method and that satisfy the Wolfe conditions in Eqs. (4), (5) and go to step (4).

Step (4): Update the variables: \( x_{k+1} = x_k + \lambda_k d_k \) and compute \( f(x_{k+1}), g_{k+1} \)

and \( s_k = x_{k+1} - x_k, \ y_k = g_{k+1} - g_k \).

Step (5): Check: if \( \|g_{k+1}\| \leq \varepsilon \) then stop. Else continue.

Step (6): The search direction: We compute the scalar \( \beta_k^{(\text{New})} \) by use the equation (27) and set \( k = k + 1 \), and go to step (4).
5. The Convergence Analysis:

Theoretical Properties for the New CG-Method.

In this section, we focus on the convergence behavior on the $\beta^{New}_k$ method with exact line searches. Hence, we make the following basic assumptions on the objective function.

**Assumption 1:**

$f$ is bounded below in the level set $L_{x_0} = \{x \in \mathbb{R}^n | f(x) \leq f(x_0) \}$; in some neighborhood $U$ of the level set $L_{x_0}$, $f$ is continuously differentiable and its gradient $\nabla f$ is Lipschitz continuous in the level set $L_{x_0}$, namely, there exists a constant $L > 0$ such that:

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \text{for all } x, y \in L_{x_0}$$

(28)

5.1 Sufficient Descent Property:

We will show that in this section the proposed algorithm that defined in the equations (27) and (3) satisfy the sufficient descent property that satisfy the convergence property.

**Theorem 1:**

The search direction $d_k$ that generated by the proposed algorithm of modified CG satisfy the descent property for all $k$, when the step size $\lambda_k$ satisfied the Wolfe conditions (4),(5).

**Proof:** we will use the indication to prove the descent property, for $k = 0$, $d_0 = -g_0 \Rightarrow d_0^T g_0 = -\|g_0\| < 0$, then we proved that the theorem is true for $k = 0$, we assume that $\|s_k\| \leq \eta ; \|g_{k+1}\| \leq \Gamma$ and $\|g_k\| \leq \eta_1$ and assume that the theorem is true for any $k$ i.e. $d_k^T g_k < 0$ or $s_k^T g_k < 0$ since $s_k = \lambda_k d_k$, now we will prove that the theorem is true for $k + 1$ then:

$$d_{k+1} = -g_{k+1} + \beta^{(New)}_k d_k$$

(29)

$$\beta^{new}_k = \left(1 - \left[\frac{y_k^T s_k}{\|s_k\|^2}\right]\right) \frac{y_k^T g_{k+1} \left(2A + 2g_{k+1}^T s_k\right)}{\|s_k\|^2}$$

(30)
i.e. \( d_{k+1} = -g_{k+1} + \left( 1 - \frac{y^T_k s^T_k}{\|s_k\|^2} \right) y^T_k g_{k+1} \left( 2A + 2g^T_{k+1}s_k \right) d_k \)  
\[(31)\]

Multiply both sides of the equation (31) by \( g^T_{k+1} \) we get:
\[g^T_{k+1} d_{k+1} = -\|g_{k+1}\|^2 + \left( 1 - \frac{y^T_k s^T_k}{\|s_k\|^2} \right) y^T_k g_{k+1} \left( 2A + 2g^T_{k+1}s_k \right) g^T_{k+1} d_k\]
\[(32)\]

\[
\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq \left( 1 - \frac{y^T_k s_k}{\|s_k\|^2} \right) y^T_k g_{k+1} \left( 2A + 2g^T_{k+1}s_k \right) \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2}
\]
\[(33)\]

\[
\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq \left( y^T_k g_{k+1} \left( 2A + 2g^T_{k+1}s_k \right) \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2} \right) \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2}
\]
\[(34)\]

\[
\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq \left( y^T_k g_{k+1} \left( 2A + 2\|g_{k+1}\|s_k \right) \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2} \right) \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2}
\]
\[(35)\]

Using strong Wolfe condition
\[\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq 2A \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2} + 2\|g_{k+1}\|s_k \frac{-\rho g^T_{k+1} d_k}{\|g_{k+1}\|^2} s_k \]
\[(37)\]

\[\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq 2A \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2} \]
\[(38)\]

Using \( S=\lambda d \)
\[\frac{g^T_{k+1} d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq \frac{2A}{\lambda} \frac{g^T_{k+1} d_k}{\|g_{k+1}\|^2} \leq 1
\]
\[(39)\]
Where $0 < \lambda < 1$

$$\frac{\|g_{k+1}\|^2}{g_k^T d_{k+1} + \|g_{k+1}\|^2} \geq \frac{\lambda}{2A} \frac{\|g_{k+1}\|d_k}{g_k^T d_k} = \delta > 1$$

(40)

$$\frac{g_k^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \leq \frac{1}{\delta}$$

(41)

$$g_k^T d_{k+1} + \|g_{k+1}\|^2 \leq \frac{1}{\delta} \|g_{k+1}\|^2$$

(42)

$$g_k^T d_{k+1} \leq -(1 - \frac{1}{\delta}) \|g_{k+1}\|^2$$

(43)

Let $c = (1 - \frac{1}{\delta})$

(44)

Then $g_k^T d_{k+1} \leq -c \|g_{k+1}\|^2$

(45)

For some positive constant $c > 0$. This condition has often used to analyze the global convergence of conjugate gradient methods with inexact line search.

5.2 Global Convergence Property:

The conclusion of the following lemma used to prove the global convergence of nonlinear conjugate gradient methods, under the general Wolfe line search.

**Lemma 1:**

Suppose assumptions (1) (i) and (ii) hold and consider any conjugate gradient method (27) and (3), where $d_k$ is a descent direction and $\lambda_k$ is obtained by the strong Wolfe line search. If

$$\sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} = \alpha$$

(46)

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

(47)
For uniformly convex functions that satisfy the above assumptions, we can prove that the norm of \( d_{k+1} \) given by (27) is bounded above. Assume that the function \( f \) is a uniformly convex function, i.e. there exists a constant \( \mu \geq 0 \) such that for all \( x, y \in S \),

\[
(g(x) - g(y))^T (x - y) \geq \mu \|x - y\|^2,
\]

(48)

Using lemma 1 the following result can be proved.

**Theorem 2:**

Suppose that the assumptions (i) and (ii) hold. Consider the algorithm (3), (27). If \( \|s_k\| \) tends to zero, and there exists nonnegative constants \( \eta_1 \) and \( \eta_2 \) such that:

\[
\|g_k\|^2 \geq \eta_1 \|s_k\|^2, \quad \|g_{k+1}\|^2 \geq \eta_2 \|s_k\|
\]

and \( f \) is a uniformly convex function, then.

\[
\lim \inf_{k \to \infty} \|g_k\| = 0
\]

(50)

**Proof:** From eq. (27) We have:

\[
\beta_{k}^{\text{new}} = \left( 1 - \left[ y_k^T s_k \right] \left[ s_k \right] \right) y_k^T g_{k+1} \left( 2A + 2g_{k+1}^T s_k \right) \left[ s_k \right]
\]

\[
|\beta_{k}^{\text{new}}| = \left| 1 - \left[ y_k^T s_k \right] \left[ s_k \right] \right| y_k^T g_{k+1} \left( 2A + 2g_{k+1}^T s_k \right) \left[ s_k \right]
\]

(51)

\[
\leq \left( 1 - \left[ y_k \left[ s_k \right] \right] \right) y_k \left[ g_{k+1} \right] \left( 2A + 2\left[ g_{k+1} \right] s_k \right) \left[ s_k \right]
\]

\[
\leq \left( y_k \left[ g_{k+1} \right] \right) \left( 2A + 2\left[ g_{k+1} \right] s_k \right) \left[ s_k \right]
\]

(53)

But \( \|y_k\| \leq L \|s_k\| \). Then
\[
\leq \left( L \| s_k \| \| g_{k+1} \| \right) \frac{(2A + 2 \| g_{k+1} \| \| s_k \|)}{\| s_k \|^2}
\] (55)

\[
\leq (L \| g_{k+1} \|) \frac{(2A + 2 \| g_{k+1} \| \| s_k \|)}{\| s_k \|}
\] (56)

\[
\leq (L \Gamma) \frac{(2A + 2 \Gamma \eta)}{\| s_k \|}
\] (57)

Hence,

\[
\| d_{k+1} \| \leq \| g_{k+1} \| + |\beta_k|^N \| s_k \|
\] (58)

\[
\| d_{k+1} \| \leq \gamma + (L \Gamma) \frac{(2A + 2 \Gamma \eta)}{\| s_k \|} = \gamma + (2AL \Gamma + 2 \Gamma^2 L \eta)
\] (59)

\[
\sum_{k \geq 1} \frac{1}{\| d_{k+1} \|^2} = \infty
\] (60)

\[
\frac{1}{\| \gamma + (2AL \Gamma + 2 \Gamma^2 L \eta) \|^2} \sum_{k \geq 1} 1 = \infty
\] (61)

6. Parallel Tangent Method:

The name of parallel tangent (PARTAN) has no significance as far as the mechanics of the search procedure are concerned; however, the name has an interesting geometrical origin, which shown in the two-dimensional case of Fig. 1. [10].
Figure 1. Locus of the search for a quadratic function.

The strong point common to all PARTAN methods, is that the acceleration step from $x_0$ through $x_2$ to $x_3$ is taken through the two points $x_0$ and $x_2$ at which the two parallel lines $L_0$ and $L_2$ are tangent to the equi-magnitude contours. To see this consider any two lines in the $x_1x_2$ plane which are parallel and which intersect a straight ravine of $f(x_1, x_2)$ (Fig. (2)). Observed that the point of tangency defines a line, which parallels the ravine. Hence, by searching along the parallel ravine-line, we effectively follow the ridge. The gradient descent searches are used to find $x_1, x_2, x_4, x_6, \ldots$ and acceleration steps are used to locate $x_3, x_5, x_7, x_9, \ldots$. With PARTAN, the acceleration steps conducted through the following pairs of points:

$$(x_0, x_2), (x_1, x_4), (x_3, x_6), \ldots, (x_{2k-3}, x_{2k}), \ldots$$

The locus of the gradient-PARTAN search would look as depicted in Fig. (3) below.

Figure 2. The path taken by the gradient-PARTAN.

6.1. A General Outlines of the PARTAN Algorithm:

Starting procedure: For the first step,

Let, $d_0 = -g_0$ and $x_1 = x_0 + \lambda d_0$. Next, choose $d_2 = -g_2$

Then, the fourth point is generated by moving in direction that is collinear with $(x_3-x_1)$ so that $d_3 = -(x_3-x_1)$

This referred to as an acceleration step. Continuing the procedure:

After determining $x_4$, the procedure continued by successively alternating gradient and acceleration steps. Thus


\[ d_i = -g_i \quad \text{for } i=0,2,...,2n-2 \quad (62) \]

\[ d_i = -(x_i - x_{i-2}) \quad \text{for } i=3,5,...,2n-1 \quad (63) \]

This method will reach the minimum of an \( n \) dimensional quadratic surface in no more than \( 2n \) steps[11].

7. Sine - Cosine Algorithm (SCA):

Mathematical sine - cosine algorithm is one of the meta-heuristic suggest by (Seyedali Marjalili 2016) which depends in general on sine and cosine functions that starts with improving a set of arbitrary solutions, then we estimate these solutions repeatedly using objective function which improved by a set of rules representing the essence of improving technique. Since techniques based on community aim at optimization for improving problems, there no guarantee to find solution in one term. With existence of a sufficient number of arbitrary solutions and improving steps (repetition), there is high probability to get optimum solutions and global values. SCA method based on finding and improving solutions, changing

\[
X^{t+1}_k = X^t_k + r_1 \times \sin(r_2) \times \left| r_3 P^t_k - X^t_k \right| 
\]

\[
X^{t+1}_k = X^t_k + r_1 \times \cos(r_2) \times \left| r_3 P^t_k - X^t_k \right| 
\]

Where sine and cosine, are the well know Mathematic functions. \( X_i^t \) is the present solution position in dimension \( i \)-th. With repetition \( t \)-th. \( r_1, r_2, r_3 \), are arbitrary numbers, as well as \( \| \) absolute value \( , r_4 \) is an arbitrary number with in period \([0,1]\). [9]

If we combined eqs. (64) and (65), we get the following:

\[
X^{t+1}_k = \begin{cases} 
X^t_k + r_1 \times \sin(r_2) \times \left| r_3 P^t_k - X^t_k \right|, & r_4 < 0.5 \\
X^t_k + r_1 \times \cos(r_2) \times \left| r_3 P^t_k - X^t_k \right|, & r_4 \geq 0.5 
\end{cases} 
\]  

(66)

Where \( r_4 \) is a random number

The range of sine and cosine in Eqs. (64) to (66) changed using:

\[
r_1 = a - t \frac{a}{T} 
\]

(67)

Where \( t \) is current iteration; \( T \) maximum number of iterations and \( a \) is a constant.

7.1. Outlines of SCA:

Step (1): Select arbitrary initial community (search agents) solutions \( X \).

Step (2): Calculate cost function for each search agents.
Step (3): Return best solution.
Step (4): Select best search agent according to cost function.
Step (5): Update $r_1, r_2, r_3$ and $r_4$.
Step (6): Update search agent position using the equation (64).
Step (7): While $t < \text{max no. iterations}$, go to step 2.
Step (8): Return best solution you got according to its degree to get global solution [9].

8. Modified Conjugate Gradient method (MCG)

It is a Hybrid method, where a conjugate factor that derived and used in Modifying conjugate Gradient algorithm CG, PT in addition to SCA method to produce on improved algorithm of high efficiency:

Below Improved method.

Step (1): Preparing and generating the initial community.
Step (2): Improving initial community by MCG, CG and PT.
Step (3): Calculating suitability function of improved community.
Step (4): Calculating the best position of all search agents to produce Improved new generation.
Step (5): Updating position of each search agents using the SCA algorithm.
Step (6): SCA works by using certain repetitions until to reach optimum value or achieve the stop condition when it finish repetition case.
Step (7): To get either minimum value of the function or about or to get the global value.

The following figure represent the new SCA-MCG algorithm
Figure (3): The Proposed SCA-MCG algorithm
9. Practical part:

To evaluate action and probability of the suggested algorithm in solving optimization problems and getting best results. It applied on a set of standard functions mentioned in table (1), to compare with SCA itself. This table includes test functions, functions extends, minimums and maximums, as well as its (F\textsubscript{min}).

| Function | Dim | Range     | F\textsubscript{min} |
|----------|-----|-----------|-----------------------|
| F\textsubscript{1}(x) = \sum_{i=1}^{n} x_i^2 | 30  | [-100,100] | 0                     |
| F\textsubscript{2}(x) = \sum_{i=1}^{n} (\sum_{j=1}^{n} x_i)^2 | 30  | [-100,100] | 0                     |
| F\textsubscript{3}(x) = \sum_{i=1}^{n} x_i^4 + \text{random}[0,1] | 30  | [-1.28,1.28] | 0                     |
| F\textsubscript{4}(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi x_i) + 10] | 30  | [-5.12,5.12] | 0                     |
| F\textsubscript{5}(x) = \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos \left(\frac{x_i}{\sqrt{i}}\right) + 1 | 30  | [-600,600] | 0                     |
| F\textsubscript{6} = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 | 2   | [-5.5]    | -1.031                |
| F\textsubscript{7}(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{|x_i|}) | 30  | [-500,500] | -418.9                |
| F\textsubscript{8}(x) = \sum_{i=1}^{n} |x_i| + \prod_{i=1}^{n} |x_i| | 30  | [-10,10] | 0                     |
| F\textsubscript{9}(x) = \max_{1 \leq i \leq n} \{ |x_i| \} | 30  | [-100,100] | 0                     |
| F\textsubscript{10} = -20\exp\left(-0.2\sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi x_i)\right) + 20 + e | 30  | [-32,32] | 0                     |

In Tables (2-4) Functions mentioned below have been applied on all mentioned algorithm. Results show the difference SCA and its improving methods, we notice that we got global values in most function which refers that improving methods for SCA, were of high efficiency, and In Tables (2-4) we notice the efficiency of the algorithms that are directly proportional to the increase in the number of search elements. The more the number of search elements increases, the better the numerical results. Note that function  F\textsubscript{10} gives constant results for all methods used and for all the number of different elements.
Table 2. compare SCA with all other Proposed Hybrid methods at No. of element =10 and iteration=500

| Function | SCA    | SCA-CG   | SCA-PT   | SCA-CG-PT | SCA-MCG |
|----------|--------|----------|----------|-----------|---------|
| F1       | 2.00E-30 | 1.46E-226 | 1.27E-227 | 0         | 0       |
| F2       | 2.13E-26 | 6.41E-210 | 1.64E-222 | 0         | 0       |
| F3       | 4.17E-78 | 0        | 0        | 0         | 0       |
| F4       | 0       | 0        | 0        | 0         | 0       |
| F5       | 0.0026593 | 0   | 0        | 0         | 0       |
| F6       | -1.0316 | -1       | -1       | 0         | 0       |
| F7       | -1257.2739 | -6.70E-18 | -4.08E-22 | -6.05E-173 | 0       |
| F8       | 7.04E-20 | 1.23E-117 | 1.22E-116 | 2.27E-215 | 0       |
| F9       | 3.71E-13 | 6.54E-111 | 2.09E-109 | 5.50E-211 | 0       |
| F10      | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 |

Table 3. compare SCA with all other Proposed Hybrid methods at No. of element =30 and iteration=500

| Function | SCA    | SCA-CG   | SCA-PT   | SCA-CG-PT | SCA-MCG |
|----------|--------|----------|----------|-----------|---------|
| F1       | 1.26E-49 | 2.70E-239 | 3.68E-238 | 0         | 0       |
| F2       | 1.72E-34 | 3.48E-225 | 8.68E-227 | 0         | 0       |
| F3       | 1.47E-80 | 0        | 0        | 0         | 0       |
| F4       | 0       | 0        | 0        | 0         | 0       |
| F5       | 0.030856 | 0        | 0        | 0         | 0       |
| F6       | -1.0316 | -1       | -1       | 0         | 0       |
| F7       | -1357.09 | -418.983 | -418.983 | -1.32E-124 | 0       |
| F8       | 1.92E-23 | 3.06E-120 | 1.21E-122 | 1.86E-220 | 0       |
| F9       | 3.66E-18 | 2.08E-115 | 2.70E-113 | 7.51E-215 | 0       |
| F10      | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 | 8.8818E-16 |
Table 4. compare SCA with all other Proposed Hybrid methods at No. of element =50 and iteration=500

| Function | SCA          | SCA-CG       | SCA-PT        | SCA-CG-PT     | SCA-MCG       |
|----------|--------------|--------------|---------------|---------------|---------------|
| F1       | 3.58E-47     | 4.04E-242    | 4.04E-242     | 0             | 0             |
| F2       | 9.74E-40     | 6.08E-229    | 6.08E-229     | 0             | 0             |
| F3       | 1.01E-88     | 0            | 0             | 0             | 0             |
| F4       | 0            | 0            | 0             | 0             | 0             |
| F5       | 0            | 0            | 0             | 0             | 0             |
| F6       | -1.0316      | -1           | -1            | 0             | 0             |
| F7       | -1401.8469   | -4.19E+02    | -4.19E+02     | -1.18E-108    | 0             |
| F8       | 1.15E-26     | 1.10E-124    | 1.10E-124     | 6.63E-225     | 0             |
| F9       | 2.48E-20     | 7.79E-119    | 7.79E-119     | 3.23E-218     | 0             |
| F10      | 8.8818E-16   | 8.8818E-16   | 8.8818E-16    | 8.8818E-16    | 8.8818E-16    |

10. Conclusions:

The process of improving and hybrid of Heuristic SCA with the suggested method of MCG and other classical ones such as CG and PT leads to increase the convergence speed and avoid falling in local solutions. It also assists in improving the resulted solution kind through the increase of detecting algorithm efficiencies. Where the results show the possibility of the improved algorithm to solve different optimization problems, after comparison, results were excellent, where global optimum value had reached in most test function. This shown in the numerical results of this study.

11. References:

[1] Al–Bayati, A.Y. and Al–Assady, N.H. (1986). "Conjugate gradient method", Technical Research, school of computer studies, Leeds University.

[2] Belazzoug, Mouhoub & Touahria, Mohamed & Nouioua, Farid & Brahimi, Mohammed. (2019). "An improved sine cosine algorithm to select features for text categorization". Journal of King Saud University – Computer and Information Sciences, pp1-11.

[3] Chandrasekaran, K. (2019)" Improved Sine Cosine Algorithm for Solving Dynamic Economic Dispatch Problem ". International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-3.

[4] Dai Y.H. and Yuan Y., "A nonlinear conjugate gradient method with a strong global convergence property", SIAM J. of Optimization, 10 (1999), 177–182.

[5] Dai, Y. H. and Zhang, H.,(2001),"An Adaptive Two-Point Step-size Gradient Method", Numerical Algorithm,27(2001) 377-385.
[6] Ekiz, Simge & Erdoğmuş, Pakize & Özgür, Büşra (2017). "Solving Constrained Optimization Problems with Sine- Cosine Algorithm". Periodicals of Engineering and Natural Sciences ISSN 2303-4521, Vol.5, No.3, November 2017, pp. 378–386.

[7] Gholizadeh1, S. & Sojoudizadeh, R. (2019). "Modified Sine-Cosine Algorithm for Sizing Optimization of Truss Structures with Discrete Design Variables". Int. J. Optim. Civil Eng., 2019; vol.(9), no.(2):pp.195-212.

[8] Kumar1, Lalit & Bharti1, Kusum Kumari. (2019). "A novel hybrid BPSO–SCA approach for feature selection". Natural Computing. Springer, published online.

[9] Mirjalili, S. (2016). "SCA: A Sine Cosine Algorithm for solving optimization problems". Knowledge Based Systems, 96, pp120–133.

[10] Mitras, B.A. and Abdul-Jabbar, S., “Pattern Recognition Using Particle Swarm Optimization with Proposed a New Conjugate Gradient Parameter in Unconstrained Optimization”, Journal of Al-Nahrain University Vol.19 (3), September, 2016, pp.138-147.

[11] Mitras, B. A., and Rasheed, K. B., (2006), “Two new approaches for PARTAN method”. Raqi journal of Statistical Science (9), pp.40-56.

[12] Ramzy, S.A. and Dunia, S., (2018). "Chaotic Sine-Cosine Optimization Algorithms". International Journal of Soft Computing 13(3):108-122.

[13] Rizk-Allah, R. M. (2017) "A Hybrid Sine Cosine Optimization algorithm for Solving Global optimization Problems". Basic Engineering sciences Dept. – Menoufia University- Egypt Scientific Research Group in Egypt (SRGE) May, 13, 20.

[14] Sindhu, R. & Ngadiran, Ruzelita & Yacob, Yasmin Mohd & Zahir, Nik Adilah Hanin & Hariharan, M. & Polat, Kemal. (2019). "A Hybrid SCA Inspired BBO for Feature Selection Problems" Hindawi Mathematical Problems in Engineering Volume 2019, Article ID 9517568, 18 pages.