International Journal of Computer Networks and Communications Security

VOL. 7, NO. 3, MARCH 2019, 49–60
Available online at: www.ijcncs.org
E-ISSN 2308-9830 (Online) / ISSN 2410-0595 (Print)

Modified Grey Wolf Optimization Algorithm by using Classical Optimization Methods

Radwan Basim Thanoon[1] and Prof. Dr. Ban Ahmed Mitras[2]

[1] M.sc. Student, department of mathematics, college of computer sciences & mathematics, Mosul

[2] Prof. Dr. department of mathematics, college of computer sciences & mathematics, Mosul

[1]bassimrathwan@gmail.com, [2]dr.banah.mitras@gmail.com

ABSTRACT

The Grey Wolf optimization algorithm represents one of a post- intuition algorithms, which was proposed first time in 2014 by Mirijalili, an algorithm based on swarms intelligence and community intelligence which inspired from the behavior of Grey wolves. It has excellent properties that exceed the characteristics of the other swarms intelligence because it is simple, flexible, easy to use and capable to developing and had a special ability to achieve the right balance between exploration and exploitation during research. Two Hybrid algorithms of the Grey wolves algorithm were proposed in this paper with two classical algorithms: (Conjugate Gradient Algorithm) and the second is (Parallel Tangent Algorithm). Each of the two algorithms above improves the elementary community randomly generated as the primary community of the Grey Wolves optimization algorithm using the characteristics of the two classical algorithms above. The test was applied to (10) high-efficiency optimization functions with different dimensions and different iteration . The results of the hybrid algorithms were excellent, encouraging and superior to the original algorithms. The new algorithms showed very high efficiency. The hybrid algorithms achieved optimal solutions by achieving the most minimum value (f mini) for most of these functions that have been tested statistically by calculating the minimum average values for more than one implementation.

Keywords: *Grey wolf optimization, post- intuition algorithms, Conjugate Gradient methods, Parallel Tangent method, unconstrained Optimization*.

1 INTRODUCTION

It is not only man who always seeks to reach everything that is ideal in this world, but everything in this universe is trying to get the optimal position possible for itself, starting with the atom, When objects or systems are formed, atoms will try to reduce the activity of their electrons as much as possible so that the body or system remains stable; Such processes are produced by purely physical laws; In biology there is a common biological principle " survival for best" , with biological evolution ;this will lead to the best adaptation of the species to the surrounding environment [1].

Optimization models try to express in mathematical terms the goal of solving the problem in a "better" way. This may mean managing a business to increase profits, increase efficiency, reduce risk. In order to solve a particular problem in an optimal manner. Therefore, models in various areas of life have been improved because human is always looking for the best [2] .

Optimization can be defined as a science to identify "the best solutions to mathematical issues and include the solution and computerization of them, and has a wide range of theoretical and practical applications in this field [3] .

However, due to the rapid development of all areas of life, the use of modern optimization approaches or non-conventional optimization methods in the approaches of solving complex engineering optimization problems in recent years has required this.

These include the modern algorithm, Simulated Annealing, Particle Swarm Optimization, Ant Colony Optimization, Fuzzy Optimization, Genetic Algorithms [4] .

More recently, global optimization has become increasingly complex and has attracted many researchers to look for ways to solve overall optimization issues. The goal is to optimize the objective function to be identified on a particular area of research [5, 6].

The methods for solving optimization issues are divided into two types of algorithms: deterministic Algorithms and Stochastic Algorithms.

Most of the classical algorithms are specific algorithms. Some of the algorithms use gradient information. They are called slope-based algorithms. For example, Newton-Raphson method [7].

Random algorithms have two types of algorithms, although the difference between them is small: heuristic algorithms and Meta-Heuristic (post- intuition) algorithms.

The Heuristic Algorithm is based on a simple principle, the result and the attempt. The search process is as follows: The solution starts with one approximate that is updated over the course of iteration such as Simulated Annealing and Hill Climbing. The meta-heuristic algorithm is used for a range of solutions (population) so that the search process begins with a random initial group (multiple solutions). This is done to enhance the population over the course of the iterations such as optimization algorithm Particle Swarm Optimization, Ant Colony Optimization, Genetic Algorithm, and so. This algorithm is characterized by qualities that are required to the work. So there are reasons that led to the development of meta-heuristic algorithms are summarized in four main reasons:

1. Simplicity: post- intuition algorithms are fairly simple, inspired by simple concepts related to physical phenomena or animal behaviors that have helped to work with computer simulation.
2. Flexibility: Flexibility refers to the applicability of the post- intuition algorithm to different problems without any particular changes in algorithm structure.
3. Derivation: Almost all meta-heuristic (post- intuition) algorithms do not contain derivation as opposed to gradient-based optimization as the process begins with the random solution and not need to calculate the derivative.
4. The ability to avoid solutions that fall within the local optimization Avoidance [8].

Two important elements of the meta-heuristic algorithm (post- intuition algorithm):

1. Exploration.
2. Exploitation.

Exploration and Exploitation are key elements of any algorithm of meta -heuristic algorithm.

Exploration is the exploration of space for a "better" solution between the set of solutions, while exploitation means verification of the exploration stage, which helps to bring about global optimization [6] .

The problem of research focused on finding the best solution for the overall optimization of unrestricted large scale issues without falling into the trap of local solutions.

The importance of the research is to improve the performance of Grey wolf optimization algorithm (CG) by combining it with two classical methods: conjugate gradient and parallel tangent. It is a proposed approach for solving high-quality unrestricted optimization issues (NP-hard).

The aim of the research is:Two hybrid algorithms based on Grey wolf optimization are proposed as follows:

1. Proposed a new hybrid algorithm consisting of Grey wolf optimization algorithm with traditional Conjugate Gradient method (GWO- CG).
2. Proposing a new hybrid algorithm consisting of the Grey wolves optimization algorithm with Parallel Tangent method (GWO- PARTAN). Both works to solve high-level issues.
3. Swarm Intelligent:

There is a rule says that “ the individual does not know thing and the group knows everything” and from this idea could have emerged the idea of (Swarm Intelligent). It can be defined as a branch of Artificial Intelligent which studies the collective behavior and characteristics of complex, self-organizing and decentralized systems with a social structure [9].

4. Grey Wolf Optimization Algorithm (GWO)

The Grey wolf optimization algorithm (GWO) is based on hunting behavior and social behavior. Nature-inspired algorithms over the past years have become very important, allowing these algorithms to be accepted and applied in different life fields because they are characterized by simplicity and flexibility. Success and related challenges in this algorithm depend on the control of parameters.

Post- intuition algorithms have extended coverage of many different areas of life such as Genetic Algorithm (GA), Optimization Ant Colony (ACO), Particle Swarm Optimization (PSO), and evolutionary programming known not only to computer scientists but also when scientists in most areas of life are different and have many applications. So there are still problems that can be solved with new optimization tools "better" than the current optimization tools. The Grey wolf algorithm (GWO) represents a mathematical model and computer simulator that simulates the driving hierarchy and the hunting state of Grey wolves in nature. The GWO algorithm was proposed by Mirijalili in 2014 [10] .

In Figure (1), Grey wolves are composed of four levels alpha (α), beta (β), delta (δ) and omega (ω)

and had particularly important in the hierarchy as shown in the following figure:

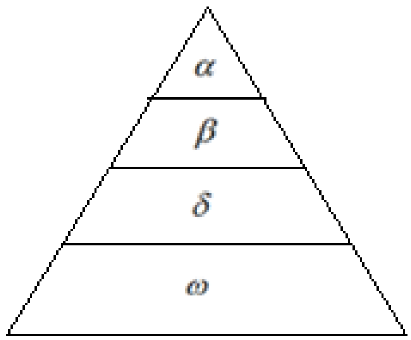

Fig. 1. Represents The Hierarchy Sequence Of Grey Wolves[11].

The wolf alpha (α) (male or female)or both is the group's leader. The whole group recognizes the wolf element (alpha) by stabilizing its tails. The alpha group is called the dominant wolf. Only alpha wolves (α) are allowed to mate in the group that Alpha (α) may be the strongest member of the group because it is the best in terms of management and responsible for deciding on hunting time, sleep time, waking up, etc., while the rest of the wolves in the group are forced to obey the orders of the wolf Alpha (α). All of its orders must be followed by the members of the group which indicates that the organization and discipline of the group represents more important than its strength [8, 12] .

The element (wolf) beta (β) (male or female) or both represents the second level after the alpha group (α) that's help the alpha group in decision-making and to understand the best alternatives to the alpha group (α) at the death or became older. The wolf is supposed to respect Beta (β), the alpha head (α), but at the same time head the wolves of the lowest level. It plays the role of advisor to the alpha element (α) of the group and beta (β) enhances alpha (α) orders in all its decisions in group [8], [12].

The groups of wolves Delta (δ) and Beta (β) obey the group of alpha wolves (α) but they are superior to a group of wolves omega (ω), which is the omega that is obedient to all upper levels. Omega plays the scapegoat [8, 12].

Search for the prey is start and then The process of conversion from one wolf to another and then the integration of wolves after the finding the prey [15].

1. Encircling of prey:

The behavior of Grey wolves to encircling around the prey was done in the mathematical modeling method of equations (1) and (2). Using these equations, the wolf updates its location within the solution area around the prey.

$$\vec{D} = \left|\vec{C} * \vec{X}_p - \vec{X}(t)\right| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}_p - \vec{A} * \vec{D} \quad (2)$$

So

t : refer to the present duplicate

$\vec{A}$ and $\vec{D}$represent transaction vectors.

$\vec{X}_p$ represents The location vector of the prey

$\vec{X}$ represent the vector of Grey wolf location .

And the vectors $\vec{A}$ and $\vec{D}$ calculated as below :

$$\vec{A} = 2 * \vec{a} * r_1 - \vec{a} \quad (3)$$

$$\vec{C} = 2 * \vec{r}_2 \quad (4)$$

The values of $\vec{a}$non-linear from 2 to 0 Over the course of iterations.

$r_1$andr_2 Represents random vectors in the standard uniform distribution in [0.1].

The update of vector for parameters (a) equation is as follows:

$$a = 2 - t(\tfrac{2}{T}) \quad (5)$$

t: refer to the present iteration

T : represent the total number of iterations [11].

The points shown in Figure (3) show that the Grey wolf updates its position in the vacuum around the prey at any random location using equations (1) and (2). The same concept can extend to all dimensions of the search area, and Grey wolves will move in excessive cubes around the best solution obtained so far as explained in the following figure [8].

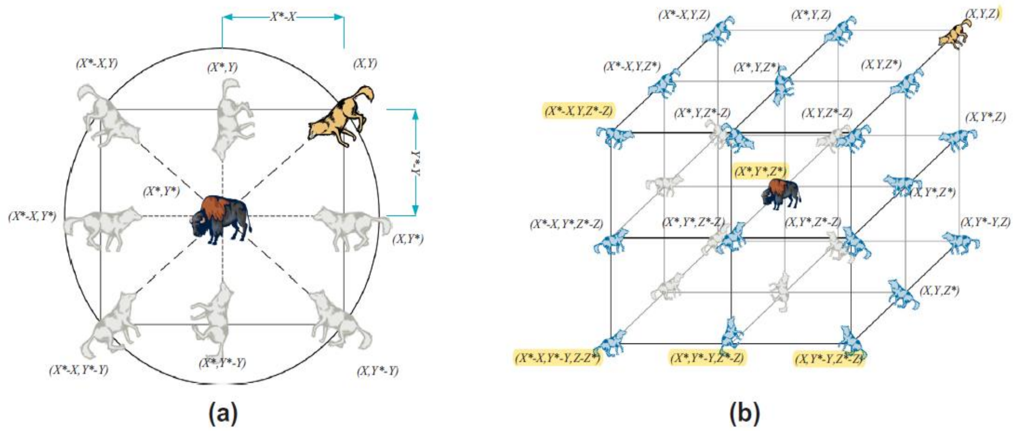

Fig. 4. Represents the 2D, 3D position and potential locations

2 HUNTING

After completing the encirclement of the prey, the Grey wolves concentrated on hunting the prey, the wolf alpha (α) usually leads the hunt, and the wolf Beta (β) and delta (δ) may share in the hunting process in the limited search area. It is not possible to know the best location (prey). When simulating the hunting behavior of Grey wolves, we assume that alpha (α) is the best initial solution and that elements or wolves beta (β) and delta (δ) have a better knowledge of the potential location of the prey The behavior of Alpha (α), beta (β) and delta (δ) can be simulated by the following equations:

$$\vec{D}_{\propto} = |\vec{C}_1 * \vec{X}_{\propto} - \vec{X}| \quad (6)$$

$$\vec{D}_{\beta} = |\vec{C}_2 * \vec{X}_{\beta} - \vec{X}| \quad (7)$$

$$\vec{D}_{\delta} = |\vec{C}_3 * \vec{X}_{\delta} - \vec{X}| \quad (8)$$

The location vector can be calculated from prey in relation to wolves Alpha (α), beta (β) and delta (δ) using the following mathematical formulas:

$$\vec{X}_1 = \vec{X}_{\propto} - \vec{A}_1 * \vec{D}_{\propto} \quad (9)$$

$$\vec{X}_2 = \vec{X}_{\beta} - \vec{A}_1 * \vec{D}_{\beta} \quad (10)$$

$$\vec{X}_3 = \vec{X}_{\delta} - \vec{A}_1 * \vec{D}_{\delta} \quad (11)$$

The best site can then be calculated from the average locations of alpha(α), beta (β) and delta (δ) as shown in the equation[8]:

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (12)$$

2.3 Grey Wolf Optimization Algorithm

The steps of the Grey Wolf Optimization Algorithm are summarized as follows:

Step 1: Randomize the population of Grey wolves.

Step 2: Randomization of Parameters a , A and C.

Step 3: Calculate the function value for each element of Grey wolves.

Step 4: For each iteration.

We find the best search element to be Xα.

We find the second best search element Xβ.

We find the third best search element Xδ.

Step 5: Since the current repeat is less than the total iterations then the search elements = Omega.

We use equations (1) and (2) of the mathematical model.

Else: Consider the search elements = Xα , Xβ , Xδ , So we use the equations from (6) to (12) of the mathematical model.

Step (6) : i=i+1

Step (7) : End

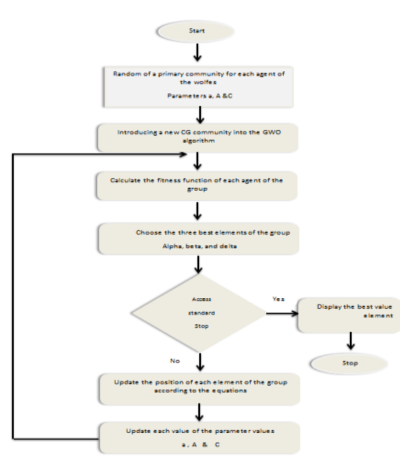

Fig. 5. Grey Wolf Optimization Algorithm

2.4 Exploration and Exploitation

Exploration and exploitation are two opposing processes. The optimization algorithm may show a specific problem in the exploration process, so the algorithm attempts to discover new parts of the problem search area by implementing sudden changes in solutions as the main goal is to discover the search space and prevent stagnation in the best localized location.

As for exploitation, the main objective is to approximating the solution to the optimal solutions that have been achieved in the exploration process by discovering the neighborhood for each solution. Therefore, gradual changes in solutions are required for optimal compatibility.

The main control of the parameter for the Grey wolves optimization algorithm (GWO) that promotes exploration is the value of variable C. This parameter always returns the random value within the period [0,2] [13] .

3 LITERATURE REVIEW

This section is devoted to reviewing the studies and researches related to the current and close research subject, Including a group of Arab and foreign studies and according to their completion time:

In 2014, both Mirjalili et al. Presented a paper in which one of the post-intuition algorithms was identified and named Grey Wolves Optimization, inspired by the behavior of Grey wolves in nature, The results were achieved through a comparative study with examples of component swarms (PSO) , It proved that the algorithm provided has good results, and this algorithm can be applied with any of the difficult questions (NP- hard) [7].

In 2016, Pradhand et al. Used an advanced method of the Grey wolves algorithm in electrical power management through an ideal operation in economic load transmission, a modern power management system [14].

In the same year, Turabieh published a paper that included the hybridization of two computerized intelligence methods, neural networks and Grey wolf examples, and used to predict heart disease. [15]

In 2017, Singh and Singh published a study that included hybridization between the swarms optimization algorithm and the Grey wolves optimization algorithm [12].

In the same year, Yassien et al. Used the Grey wolf optimization algorithm and the aggregation algorithm to find the best solution to the problem of backpack. [16].

In 2017, Kaur et al. Presented a working paper that included a parallel function scheduler using the genetic algorithm and the use of the Grey wolf optimization algorithm to reduce spare time and waiting time[17].

In the same year, Joshi and Atore presented research on how to enhance the best hunting mechanism in the Grey Wolves optimization algorithm, which focuses on enhancing the balance between exploration and exploitation in the performance of the algorithm.[6]

Also , in 2017, Jeet published research that included the use of the Grey wolves algorithm for software uniform[18]

In 2018, Faris et al. Published a paper reviewing preliminary information on optimization of Grey wolves by clarifying the context of natural formation, major processes and scientific applications. [13]

In 2018 Emary et al. Proposed a new type of Grey wolves optimization algorithm to promote the principles and properties of learning in artificial neural networks to improve their performance. [19]

As well as in 2018 Jian and Zhany used the algorithm of Grey wolves optimization to deal with the scheduling issue [20]

In the same year, Yany and Liu proposed a multi-purpose Grey wolves optimization algorithm Multi-purpose to solve the issue of tabular scheduling [21]

Kim and others in 2018 also used the Grey wolves optimization algorithm to obtain optimum coordination in the direction of the excessive electric current [9]

3.1 Conjugate Gradient Methods

Attention was paid to the associated gradient methods for two reasons, The first is that these methods are among the oldest and best known techniques for solving systems of linear equations of large dimensions. The second reason is that these methods can be adapted to solve problems of nonlinear optimization [22].

These methods have advantages that place them between the steepest descent method and the Newton method, Because these methods only require calculation of the first derivatives, they do not need to calculate and store the second derivatives needed by the Newton method, and they are faster than the steepest descent method, that is, they overcame the slow convergence of this method and because they do not need to calculate the hessian matrix or any of its approximations They are widely used to solve optimization issues of high dimension[3] [22][23]:

Conjugate gradient (CG) methods represent an important class of unconstrained optimization algorithm. The main advantages of the CG methods are its low memory requirements, its convergence speed and its poses a quadratic termination property in which the method is able to locate the minimize of quadratic function in a known finite number of iterations.

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:

$$x_{k+1} = x_k + \lambda_k d_k , \tag{13}$$

where the positive step size $\lambda_k > 0$ is obtained by a line search, and the directions d_k are generated as:

$$d_{k+1} = -g_{k+1} + \beta_k d_k , \tag{14}$$

where $d_0 = -g_0$, the value of β_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 are often based on some version of the Wolfe conditions. The standard Wolfe conditions:

$$f(x_k + \lambda_k d_k) - f(x_k) \leq \rho \lambda_k g_k^T d_k , \tag{15}$$

$$g(x_k + \lambda_k d_k)^T d_k \geq \sigma g_k^T d_k , \tag{16}$$

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

$$\beta_k^{(HS)} = \frac{y_k^T g_{k+1}}{y_k^T d_k} (Hestenes and Stiefel) \tag{17}$$

$$\beta_k^{(FR)} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} (Fletcher and Reeves) \tag{18}$$

$$\beta_k^{(PRP)} = \frac{y_k^T g_{k+1}}{g_k^T g_k} (Polak - Ribiere ; and Polyak) \tag{19}$$

$$\beta_k^{(CD)} = -\frac{g_{k+1}^T g_{k+1}}{g_k^T d_k} (Conjugate descent) \tag{20}$$

$$\beta_k^{(LS)} = -\frac{y_k^T g_{k+1}}{g_k^T d_k} (Liu and Stoery) \tag{21}$$

Outlines of the CG methods [26]:

Step1: Given $x_0 \in R^n$, compute $g_0=\nabla f(x_0)$. If $\|g_0\| \leq \varepsilon = 10^{-6}$ then
stop, Otherwise, set k=1 and continue

Step 2: Set $d_k = -g_k$.

Step 3: Compute $x_{k+1} = x_k + \alpha_k d_k$.

Step 4: If $\|g_{k+1}\| < \varepsilon$ then stop else continue.

Step5: Compute the new search direction defined by:

$$d_{k+1} = -g_{k+1} + \beta_k d_k$$

Where β_k is the conjugancy coefficient.

Step6: Check for convergence if $\|g_{k+1}\| < \varepsilon$ then stop, otherwise set
k=k+1and go to step 3.

3.2 Parallel Tangent Gradient (Zigzagging Phenomenon):

The descent algorithm tends to be near the optimal point as small orthogonal steps are taken, which is known as the Zigzagging phenomenon. To understand the Zigzagging phenomenon, let us look at an objective function with concentric ellipse lines with the minimum limit for the best point of P * as shown in Figure 6 below [6]:

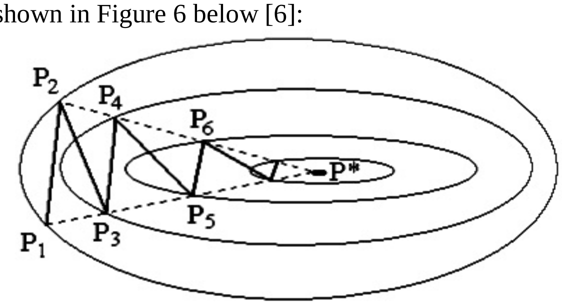

Fig. 6. The Zigzagging phenomenon in the ellipse

3.2.1 Parallel Tangent Gradient: Parallel Tangent Gradient: Parallel Tangent is a technique that combines many simple characteristics of gradient methods. It uses the geometrical property in the formation of quadratic functions. It works in the form of elliptical lines. It has several formulas, including the gradient of the parallel Tangent , which is multidimensional in the combined gradient methods. Therefore, this technique is based on the improvement of the steep way. In Figure (4) we can see two ascending rectangles approaching the optimal point P*. This indicates that the search is from point P_3 and does not lead to the direction of point P_4. but along the straight line connecting P_1 and P_3[23]. The procedure to be achieved in the objective function to the minimum on successive straight lines is to determine the direction alternating position by the previous point or gradient directions. This method does not involve the construction of a reciprocal light with a conjugate vector that can be created through two direction vector with the conjugate. The property is based on convergence and is called the parallel Tangent method [24].

In Fig. 6 this way the optimal point P* can be located after three steps as:

1. From P_1 to P_2 along the gradient to P_1.
2. From P_2 to P_3 along the gradient to P_2.
3. At last from P_3 to P_1 along the gradient to P_3.[23]

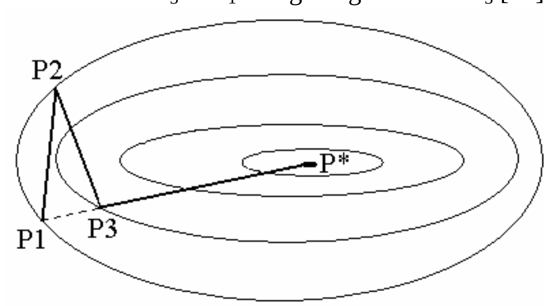

Fig. 7. Represents the base of the parallel tangent gradient in finding the optimal point P*

Figure 7 shows a diagram of the parallel tangent gradient algorithm. Note that the dots were numbered. When individual points, for example, after the point P_3, mean (p_5 , p_7 , p_9 …..) It is the result of gradation in the search, while the post even P_2 points mean (p_4 , p_6 , p_8 …..) acceleration can be obtained, in other words using a point difference system, be either odd or even. The P_{2k} point of the conjugation points can be determined by acceleration from P_{2k-4} to P_{2k-1} when K=2,3,…..,N that means

$P_{2k}= \Omega(P_{2k-1} , P_{2k-4}) , K \geq 2$

The Ω is an acceleration function, so acceleration requires a process of taking the minimum point on the line that falls between the odd points and even points.

Consider that the gradient of the parallel tangent shown in Figure (7) reduces different functions in the number of variables. P_0 is the starting point, then P_2 and P_3 are standard gradually steps, and an optimal duplicates is followed for the two steps. In the first step, the acceleration and in the second step is the standard gradient [25].

Figure (8): The path taken by the gradient-PARTAN.

3.2.2 Parallel Tangent Algorithm

1. Choose a starting point and assume $d_0= - g_0$

So $x_1 = x_0 + \lambda d_0$

after this choosing $d_2 = - g_2$

Then, the fourth point is created by moving in one direction with ($x_3 – x_1$).

$d_3= - (x_3 – x_1)$

This point is referred to as an acceleration step after the determination of whether to follow the process in succession with gradual rotation and acceleration of steps. And so on

2. $d_i = -g_i$ for i = 0,2, ... ,2n - 2

$d_i = - (x_i - x_{i-2})$for i = 3,5, ... ,2n - 1

In this way we will reach the (minimum) in (n) of the square surfaces dimensions not exceeding 2n steps.

In general the gradient is interrelated and not reciprocal and follows the following characteristics :

1. The direction of the search is most appropriate $d_i^T g_i < 0$
2. The vectors $(x_2 - x_0), (x_2 - x_0), \ldots, (x_{2n} - x_{2n-2})$ mutually conjugancy.
3. Points $x_4, x_6, \ldots, x_{2n}$,The minimum search area is a straight extend by $d_1 \& d_2$, $g_2 \& g_4, \ldots (d_1, d_2, \ldots, d_{2n})$
4. Gradient vectors $g_2, g_2, \ldots, g_{2n}$be orthogonal.

The parallel Tangent algorithm stops when $\| g_{k+1} \|$ is small enough.

The ideal calculation should end in n of iterations , We choose an elementary point x_0, especially when the algorithm approaches k <n of iterations if the Hessian Matrix and only k function had special values. These properties are tracked because of repetitive relationships in d_i which are designed to ensure that the search direction is associated with the Hessian Matrix. The behavior of the parallel tangent algorithm depends on the exact accuracy of the calculation over the accuracy of these numerical results [24].

4 HYBRID GREY WOLF ALGORITHMS

That the single evolutionary algorithm can give good solutions to many issues, but there are several types of issues that the evolutionary algorithm can fail to obtain a suitable solution (optional) for it and this paves the way for hybridization of evolutionary algorithms with each other or with other optimization algorithms or with classical algorithms, Therefore, many possible causes of hybridization can be described as follows:

1. To improve the performance of the evolutionary algorithm (eg increase the speed of convergence).
2. Improve the quality of solutions obtained from the evolutionary algorithm.
3. Make the evolutionary algorithm part of a larger system.

This is why many researchers have over the past years combined the algorithms of ordinary or difficult problem (NP - hard). The best results have been found for many optimization problem via hybrid algorithms. Therefore, the essence of the hybridization process for general and diverse solutions is its ability to deal with many real world problems.

This aspect deals with the proposal of two types of hybrid algorithms, namely hybridization of the meta-heuristic algorithm with classical algorithms, The Grey wolf Optimization algorithm (GWOA) was linked With the Conjugate Gradient Algorithm (CG) in the first proposed algorithm. In the second algorithm, the algorithm of Wolf Optimization Algorithm with parallel tangent algorithm(PARTAN).

5 A NEW HYBRID GWO WITH CG ALGORITHM

In this section, a new way to solve optimization problems was proposed, a proposed hybrid algorithm by linking the evolutionary algorithm ideas to GWO with the traditional optimization concepts Conjugate Gradient Algorithm named (GWO-CG) In this algorithm the process is divided into each iteration to two stages, in the first stage the randomization and speed algorithm is used GWO, and in the second phase the FR- CG algorithm is used. The Grey wolves algorithm in its properties is distinguished from other meta-heuristic algorithms in the search for the best three wolves, Alpha (α) and Beta (β) and delta (δ). These are the three best solutions. The maximum benefit of this property was achieved when hybridization was carried out. The steps of the hybrid algorithm (GWO- CG) as follows:

Step 1: Generate initial population by creating a primary community, and creating the parameters a ,A ,C.

Step 2: Enter random values based on CG.

Step 3: Calculate the fitness function to Each search element represents the individual wolf distance from the prey.

Step 4: Calculate the top three locations in the search elements of alpha (α), beta (β) and delta (δ). This property can produce a new generation of children.

Step 5: Update the location of each search element using the properties of the algorithm: searching for the prey, encircling the prey, hunting and attacking the prey.

Step 6: Update the new generation location using equations (6), (7) and (8).

Step 7: The classical algorithm CG of parents and children is introduced together in step (5) as an initial population algorithm for Grey wolves optimization. It is an initial society for the Grey wolves algorithm. The fitness function is then calculated for this population as soon as the process of improvement begins. After the best solution is obtained, GWO algorithm performs a number of iterative steps until the stop condition is met. (Stops execution at maximum iterations).

Step 8: The improved community is arranged based on the three best locations: Alpha (α), Beta (β) and Delta (δ).

Step 9: After the improved population order comes the role of the third characteristic of Grey wolves algorithm (attacking and hunting). To reach the maximum allowed level, the low-fitness elements are removed and the process is repeated until the optimal solution is reached and the number of iterations is completed.

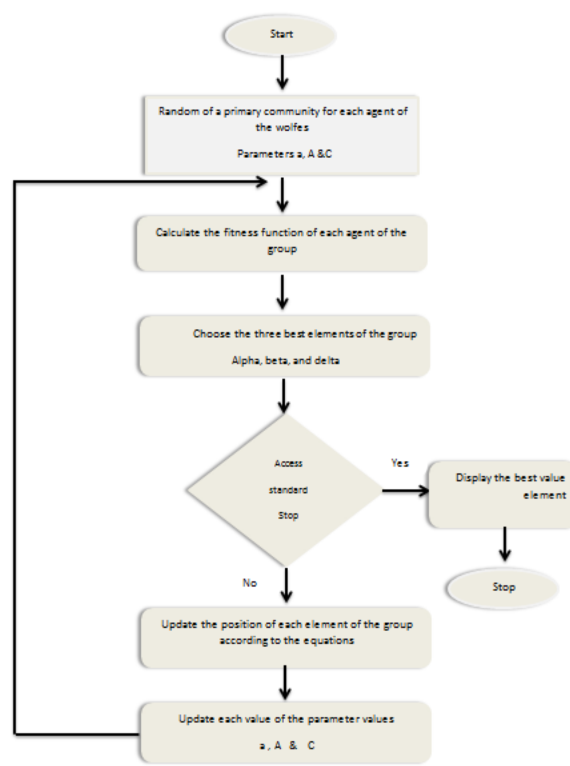

Fig. 9. Hybrid Grey Wolf Optimization Algorithm with CG(GWO-CG)

6 A NEW 2 HYBRID GWO ALGORITHM WITH PARTAN METHOD

In this section, a new hybrid algorithm was proposed by associating an algorithm of swarms intelligence with a classical algorithm called the Parallel Tangent Algorithm (GWO-PARTAN) , In this algorithm, the process is divided into two stages. In the first stage, the GWO algorithm is used. The second phase uses the PARTAN algorithm, which is divided into three sections. It is a conjugate gradient that uses the value of the beta HS, The difference of the derivative in the case if the point is even, the difference points ((x_i-x_(i-1) if the points is odd .

The important characteristic of meta-heuristic algorithms is to find the optimal solution without the need to calculate the derivatives, while in traditional methods, especially (PARTAN method), there is a need to calculate the derivative because the solution moves in the negative direction (d_k= - g_k).

7 PRACTICAL SIDE

Purpose To evaluate the performance of the proposed algorithms in solving optimization problems , the first suggested algorithm was tested

GWO-CG and the second proposed algorithm GWO-PATRAN, Using 10 standard functions to compare with the same Grey wolves algorithm. The table (1) shows the details of the test functions, the special range, the minimum value (f min), the lower and upper limits of each function, the number of iterations used (500) iterations , As for the number of search elements, in the table (2) there are (5) elements and (10) elements in table (3) and (15) elements in the table (4) .

Applying the test using a laptop that has the following specifications: CPU Processor speed is 2.4GHZ and memory RAM is 3GB and using MATLAB program R2013a work on Windows 10 operating system.

Table1: Test Function

Function	Dim	Range	*Fmin*
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[−100,100]	0
$F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[−10,10]	0
$F_3(x) = \sum_{i=1}^{n} (\sum_{j-1}^{i} x_i)^2$	30	[−100,100]	0
$F_4(x) = \max_i\{\lvert x_i \rvert . 1 \le i \le n\}$	30	[−100,100]	0
$F_5(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{\lvert x_i \rvert})$	30	[−500,500]	−418.9829
$F_6(x) = \sum_{i=1}^{n} \lvert x_i^2 - 10\cos(2\pi x_i) + 10 \rvert$	30	[−512,512]	0
$F_7(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right) + 20 + e$	30	[−30,30]	0
$F_8(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[−50,50]	0
$F_9(x) = \sum_{i=1}^{n} i x_i^4 + random[0,1)$	30	[−1.28,1.28]	0
$F_{10}(x) = 4x_i^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	[−5,5]	−1.0316

Table 2: Results comparison table between GWO and new hybrid GWO - CG and New 2 hybrid GWO-PATRAN using the number of elements consisting of 5 elements and the number of iterations 500

		PARTAN			
Function	GWO	CG	S_k	y_k	GWO-CG
F1	2.2328e-09	3.0832e-27	4.6379e-27	4.3790e-27	1.1892e-45
F2	2.3594e-06	3.2192e-15	4.2835e-15	4.0841e-15	2.2985e-24
F3	1.8681	1.6359e-17	1.1352e-17	1.7053e-17	3.4996e-35
F4	0.0276	4.2924e-11	2.9231e-11	5.1530e-11	5.7062e-20
F5	-4.7268e+03	-3.9453	-3.9453	-3.9453	-3.9453
F6	5.2374	0	0	0	0
F7	3.1464e-06	1.0036e-13	6.9574e-14	7.1942e-14	8.8818e-16
F8	0.0569	0	0	0	0
F9	4.3410e-07	3.9968e-14	3.9968e-14	3.6415e-14	8.8818e-16
F10	-1.0316	-1	-1	-1	-1

Table3: Results comparison table between GWO and new hybrid GWO - CG and New 2 hybrid GWO-PATRAN using the number of elements consisting of 10 elements and the number of iterations 500

		PARTAN			
Function	GWO	CG	S_k	y_k	GWO-CG
F1	7.6587e-16	8.5109e-34	2.0088e-33	9.1180e-34	2.6257e-51
F2	6.5792e-10	4.1588e-19	1.3520e-18	4.3693e-19	4.4165e-28
F3	0.3721	3.6025e-19	1.1109e-19	2.5816e-19	1.7901e-37
F4	0.0015	6.1239e-13	4.2761e-13	8.4179e-13	4.3152e-21
F5	-5.6716e+03	-3.9453	-3.9453	-3.9453	-3.9453
F6	6.9465	0	0	0	0
F7	1.8152e-09	3.2863e-14	3.1678e-14	2.8126e-14	8.8818e-16
F8	1.9614e-15	0	0	0	0
F9	1.3534e-10	2.2204e-14	2.2204e-14	2.2204e-14	8.8818e-16
F10	-1.0316	-1	-1	-1	-1

Table 4: Results comparison table between GWO and new hybrid GWO - CG and New 2 hybrid GWO-PATRAN using the number of elements consisting of 15 elements and the number of iterations 500

		PARTAN			
Function	GWO	CG	S_k	y_k	GWO-CG
F1	1.9157e-19	4.2593e-38	7.8911e-38	6.0827e-38	3.6984e-56
F2	3.0029e-12	2.7270e-21	1.0795e-21	2.3245e-21	1.3645e-30
F3	9.2709e-04	1.3756e-21	5.6984e-21	1.7475e-21	3.7101e-40
F4	1.0599e-04	2.0872e-13	2.5048e-14	4.4127e-14	1.0510e-22
F5	-5.9941e+03	-3.9453	-3.9453	-3.9453	-3.9453
F6	8.2833	0	0	0	0
F7	3.6930e-11	2.3389e-14	2.5757e-14	2.4573e-14	8.8818e-16
F8	0.0057	0	0	0	0
F9	2.0484e-12	1.5099e-14	1.1546e-14	1.5099e-14	8.8818e-16
F10	-1.0316	-1	-1	-1	-1

Table 5: Results comparison table between GWO and new hybrid GWO - CG and New 2 hybrid GWO-PATRAN using the number of elements consisting of 20 elements and the number of iterations 500

		PARTAN			
Function	GWO	CG	S_k	y_k	GWO-CG
	1.9951e-23	2.0750e-41	7.4915e-41	3.0265e-41	1.5439e-59
F2	4.0263e-14	1.9742e-23	2.5578e-23	2.2690e-23	3.0354e-32
F3	8.7965e-04	3.8635e-23	7.5274e-21	7.3286e-22	6.4629e-41
F4	2.2052e-05	4.9244e-15	9.2817e-15	8.1538e-15	2.0871e-23
F5	-4.7155e+03	-3.9453	-3.9453	-3.9453	-3.9453
F6	1.0232e-12	0	0	0	0
F7	7.0434e-13	1.9836e-14	1.8652e-14	2.2204e-14	8.8818e-16
F8	0	0	0	0	0
F9	1.6905e-13	1.5099e-14	1.5099e-14	1.7467e-14	8.8818e-16
F10	-1.0316	-1	-1	-1	-1

8 DISCUSSION OF NUMERICAL RESULTS

The results in the tables above show the success of the new hybrid algorithm GWO-CG and the second hybrid algorithm GWO-PARTAN to find the optimal solution for 10 of the high-standard standard test functions compared with the original GWO optimization algorithm, which confirms the success of the hybridization process.

The proposed algorithm is the first GWO-CG and the second proposed GWO - PATRAN gave better results than the original GWO algorithm note that the functions (F6 and F8) have given the best solution is the optimal (0) The functions from (F1 to F5 and F7 and F9) We note little improvement.

9 CONCLUSIONS

The hybridization of the meta-heuristic algorithm with classical algorithms contributed to improving the performance of original meta-heuristic algorithms by increasing the speed of convergence, It also improved the quality of the resulting solutions by increasing their exploratory and exploitative capabilities. Numerical results showed the ability of hybrid algorithms to solve various optimization problems. The results of the GWO-CG and GWO-PARTAN algorithm were compared with the Grey wolves optimization algorithm itself, resulting in excellent results. The optimal overall solution was obtained for most test functions.

10 REFERENCES

[1] T. Weise, Global optimization algorithms-theory and application, Self-published 2 (2009).
[2] I. Griva, S. G. Nash and A. Sofer, Linear and nonlinear optimization, vol. 108, Siam2009.
[3] W. Sun and Y.-X. Yuan, Optimization theory and methods: Nonlinear programming, vol. 1, Springer Science & Business Media2006.
[4] S. S. Rao, Engineering optimization: Theory and practice, John Wiley & Sons2009.
[5] S. Zhang, Q. Luo and Y. Zhou, Hybrid grey wolf optimizer using elite opposition-based learning strategy and simplex method, International Journal of Computational Intelligence and Applications 16 (2017), no. 02, 1750012.
[6] H. Joshi and S. Arora, Enhanced grey wolf optimization algorithm for global optimization, Fundamenta Informaticae 153 (2017), no. 3, 235-264.
[7] S. Mirjalili, S. M. Mirjalili and A. Lewis, Grey wolf optimizer, Advances in engineering software 69 (2014), 46-61.
[8] M. Karimi and S. M. Babamir, Qos-aware web service composition using gray wolf optimizer, International Journal of Information & Communication Technology Research 9 (2017), no. 1, 9-16.
[9] C.-H. Kim, T. Khurshaid, A. Wadood, S. G. Farkoush and S.-B. Rhee, Gray wolf optimizer for the optimal coordination of directional overcurrent relay, Journal of Electrical Engineering & Technology 13 (2018), no. 3, 1043-1051.
[10] L. Korayem, M. Khorsid and S. Kassem, Using grey wolf algorithm to solve the capacitated vehicle routing problem, IOP conference series: materials science and engineering, IOP Publishing, 2015, p. 012014.
[11] Y. Ren, T. Ye, M. Huang and S. Feng, Gray wolf optimization algorithm for multi-constraints second-order stochastic dominance portfolio optimization, Algorithms 11 (2018), no. 5, 72.
[12] N. Singh and S. Singh, Hybrid algorithm of particle swarm optimization and grey wolf optimizer for improving convergence performance, Journal of Applied Mathematics 2017 (2017).
[13] H. Faris, I. Aljarah, M. A. Al-Betar and S. Mirjalili, Grey wolf optimizer: A review of recent variants and applications, Neural computing and applications (2018), 1-23.
[14] M. Pradhan, P. K. Roy and T. Pal, Grey wolf optimization applied to economic load dispatch problems, International Journal of Electrical Power & Energy Systems 83 (2016), 325-334.
[15] H. Turabieh, A hybrid ann-gwo algorithm for prediction of heart disease, American Journal of Operations Research 6 (2016), no. 02, 136.
[16] E. Yassien, R. Masadeh, A. Alzaqebah and A. Shaheen, Grey wolf optimization applied to the 0/1 knapsack problem, International Journal of Computer Applications 169 (2017), no. 5.
[17] S. Kaur, K. Kaur and A. Chhabra, Parallel job scheduling using grey wolf optimization algorithm for heterogeneous multi-cluster environment, (2017).
[18] K. Jeet, Grey wolf algorithm for software organization, Indian J. Sci. Res 7 (2017), no. 2, 214-217.
[19] E. Emary, H. M. Zawbaa and C. Grosan, Experienced gray wolf optimization through reinforcement learning and neural networks, IEEE transactions on neural networks and learning systems 29 (2018), no. 3, 681-694.
[20] T. Jiang and C. Zhang, Application of grey wolf optimization for solving combinatorial

problems: Job shop and flexible job shop scheduling cases, IEEE Access (2018).
[21] Z. Yang and C. Liu, A hybrid multi-objective gray wolf optimization algorithm for a fuzzy blocking flow shop scheduling problem, Advances in Mechanical Engineering 10 (2018), no. 3, 1687814018765535.
[22] J. Nocedal and S. J. Wright, "Numerical optimization 2nd," Springer2006.
[23] P. Moallem, S. A. Monadjemi, B. Mirzaeian and M. Ashourian, A novel fast backpropagation learning algorithm using parallel tangent and heuristic line search, Proceedings of the 10th WSEAS international conference on Computers, World Scientific and Engineering Academy and Society (WSEAS), 2006, pp. 634-639.
[24] B. A. Metras and K. B. Rasheed, Two new approaches for partan method.
[25] P. Moallem, S. A. Monadjemi and B. Mirzaeian, Ptgmvs: Parallel tangent gradient with modified variable steps for a reliable and fast mlp neural networks learning, KUWAIT JOURNAL OF SCIENCE AND ENGINEERING 35 (2008), no. 1B, 65.