

# A Collective Intelligence Strategy for Enhancing Population-based Optimization Algorithms

Azam Asilian Bidgoli\*, Shahryar Rahnamayan, SMIEEE  
Nature Inspired Computational Intelligence (NICI) Lab  
Department of Electrical, Computer, and Software Engineering  
Ontario Tech University, Oshawa, Canada  
azam.asilianbidgoli@uoit.ca, shahryar.rahnamayan@uoit.ca

**Abstract**—Population-based algorithms are a well-established category of metaheuristic optimization algorithms in which individuals collaborate with each other to find the optimal solution in a search space. During the search process, each individual provides a partial intelligence which can assist the population movement toward promising regions. In this paper, a dimension-wise strategy is proposed to collect the intelligence of whole population to generate a new trial candidate solution. For new individual, the value of each variable is calculated using the votes of a more-crowded cluster of individuals obtained on each dimension (one-dimensional clustering). Accordingly, a group of candidate solutions in the population collaborate to determine a variable value of new individual. Utilizing this strategy, collective intelligence (CI) aims the algorithm to find better candidate solutions. Since the proposed method keeps untouched all other parts of the algorithm, it can be used with any population-based algorithm. This paper presents the modification of two well-known population-based algorithms based on the proposed strategy in utilizing Collective Intelligence (CI), Differential Evolution (CIDE) and Particle Swarm Optimization (CIPSO). In conducted experiments, two proposed algorithms are compared with classical version of DE and PSO on 30 functions of CEC-2017 benchmark. The results indicate that the proposed method generates an individual with better objective function value than many of the individuals in the population which leads totally better results in overall.

**Index Terms**—Collective Intelligence, Population-based Algorithms, Differential Evolution (DE), Particle Swarm Optimization (PSO), Metaheuristics, Voting, Dimension-wise Clustering

## I. INTRODUCTION

Population-based metaheuristic algorithms are one of the well-known family of algorithms for solving the global optimization problems [1], [2]. These algorithms are often biological or social inspired algorithms which represent a range of universal problem solvers. They begin the optimization search by using an initial population which consists of candidate solutions; then they try to find the optimal solutions using an iterative and stochastic processes. Despite the importance of optimization task and designing numerous algorithms, the capability of solving many global optimization problems is still not satisfactory. Hence, researchers try to utilize various strategy to improve the existing algorithm or design new algorithms.

Two most popular families of population-based algorithms are evolutionary algorithms (EAs) [3] and swarm intelligence

algorithms [4]. In both families of algorithms, the use of some sort of implicit or explicit collective intelligence of individuals in the population can be recognized. Collective intelligence is an integration process of collective behaviors of individuals in social groups or collective functions of components in computational intelligent systems [5]. A transdisciplinary study on metaheuristic computing and social psychology may explain a set of key mechanisms of collected intelligence.

Evolutionary algorithms work based on generating new individuals using crossover and mutation operations during the optimization process. Then, elite solutions are selected to be transferred to the next generation [6]. Regarding evolutionary algorithms, one of the operations which is defined as use of the collective intelligence of the population is crossover. This operation generates new offspring using combining two or more individuals of the population. When the number of parents increases, in fact the participants on collective intelligence increase. Based on some research investigations, increasing the number of parents in the Genetic Algorithm (GA) [7] improves the convergence rate of the classical GA [8]. In all multi-parent crossovers, the parents number plays a key role to keep high the population diversity and to reduce the risk of premature convergence [9]. In [10], authors proposed taking advantage of the collective information from the  $m$  best candidate solutions in continuous DE by combining them to use the combinational solution crossover and mutation operations. In [5], the properties of collective intelligence and its applications in metaheuristic computing are represented.

Swarm-based algorithms, second group of metaheuristics, are made up of simple particles which cooperate with each other and interact with their environment to move toward the region of the optimal solution [11]. The movement of each particle is updated based on some rules to improve its behavior in terms of getting closer to the target. Since the position of the best particle is a part of rules' definition, it can be stated that the collective intelligence of population affects movement of whole population. However, in the classical version of the algorithm, the intelligence doesn't come from the whole population, instead it is from the best members of the population or the top best particles. In [12], the global best solution of the PSO was modified by taking into account all local optimums of particles multiplied by their objective function

values. As another example for collaborating of individuals of the population to find the optimal solution in the search space is Ant Colony Algorithm (ACO) [13]. The main principle founding ant-based algorithms is to leave traceable signs for letting other to follow individuals guided through them [14]. However, other algorithms such as Particle Swarm intelligence Optimization (PSO) [15] and Differential Evolution (DE) [16] use the implicit memory for their collaboration; but ACO uses an explicit memory.

All mentioned perspectives of collective intelligence in metaheuristic algorithms involve only a part of the population or use the its benefit implicitly in their operations. There has been conducted a study on utilizing whole population intelligence but on discrete optimization problems. The voting operator has been used as collective intelligence on discrete problems in [17]. The authors generate a new individual by voting on variables' value of individuals. Increasing the participation of individuals to utilize the intelligence of more components in any type of optimization problem is the main idea of this paper.

The proposed method generates a new individual using the collective intelligence of whole population. Each variable of the new vector is obtained using a continuous voting scheme on a crowded cluster of individuals which is earned by applying a clustering algorithm. By considering this idea, whole population contribute to generate a high quality candidate solution. Because the clustering task is dimension-wise, it doesn't suffer increasing the number of dimensions. Two popular metaheuristic algorithms, DE and PSO algorithms are modified based on the proposed scheme. In order to evaluate the performance of proposed method, some experiments are conducted on functions of CEC-2017 benchmark set to compare the collective intelligence-based versions of DE and PSO (called CIDE and CIPSO, respectively) with their parents.

The remaining sections of this paper are organized as the follows. Section 2 presents a background review on applied algorithms. Section 3 provides a detailed explanations about the proposed method. The performance of the proposed method is investigated over 30 benchmark functions in the Section 4. The paper is concluded in the Section 5.

## II. BACKGROUND REVIEW

In this section, a brief explanation of parent algorithms which the proposed method are compared with, is provided. DE and PSO are two well-known population-based optimization algorithms which are described. In addition, the k-means algorithm as utilized clustering algorithm for collecting the intelligence is explained in this section.

### A. Differential Evolution

DE is an evolutionary algorithm which is originally designed to solve continuous optimization problems. The main part of DE is its mutation operator, otherwise, all other steps of algorithm is kept untouched rather than other evolutionary algorithms. Similar to other population-based algorithm, DE starts with a uniform random initialization of the population.

Objective function is calculated for each individual in the population. Next generation should be constructed using selecting the best individuals from current and produced individuals. In order to generate new individuals, an evolutionary mutation is conducted. For each individual, an offspring is generated using three randomly selected individuals from population based on Eq. 1.

$$v_{j,i} = x_{j,i_1} + F \cdot (x_{j,i_2} - x_{j,i_3}), \quad (1)$$

where  $x_{j,i_1}$ ,  $x_{j,i_2}$ , and  $x_{j,i_3}$  are three randomly selected individuals. For each variable of new individual ( $v_i$ ), the value of variables of two selected individuals are subtracted using a mutation factor,  $F$ , then the result is added to the third selected individual. This operator generates new trial candidate solution to get away from the  $x_{i_3}$  candidate and move toward the  $x_{i_1}$  and  $x_{i_2}$  candidate solutions. Some or all genes of parent are selected based on Crossover Rate ( $C_R$ ) to be replaced by genes of generated individual. After producing offsprings in number of population size, DE selects best individuals using comparing parent and its corresponding offspring. Each of them with better objective value is remained in the population to be transferred to the next generation. This steps continue to meet a stopping criterion.

### B. Particle Swarm Optimization

PSO is one of the well-known metaheuristic algorithm. It works based on movement of swarm of individuals which called particles. Initial population is generated uniform randomly. Each particle is defined using a position which means the value of variables and a velocity of their movement. Optimization forces all particles to move toward a global optimum solution. At each generation, PSO updates the position and velocity of particles based on Eq.s 2 and 3.

$$v_{id} = w \cdot v_{id} + c_1 \cdot r_1 \cdot (p_{id} - x_{id}) + c_2 \cdot r_2 \cdot (p_{gd} - x_{id}) \quad (2)$$

$$x_{id} = x_{id} + v_{id}, \quad (3)$$

where the position of a particle,  $x_{id}$ , is updated by  $v_{id}$  which shows the particle velocity. The position of the best candidate solution in a predefined neighborhood ( $p_{gd}$ ) along with the position of the best movement achieved so far associated with each particle ( $p_{id}$ ) changes the particle's velocity toward the global and local best positions.  $r_1$  and  $r_2$  are two uniform random vectors in range  $(0, 1)$ .  $c_1$ ,  $c_2$  are two parameters representing the particle's confidence in itself and in the swarm experiences, respectively and  $w$  determines the effect of previous velocity, called inertia weight. There are variant kinds of topology schemes [18] which are defined for PSO to identify  $p_{gd}$ . The most commonly used topology is a 2-neighborhood undirected regular ring topology [19]. In the ring topology, each particle,  $i$ , has two neighbors,  $i - 1$  and  $i + 1$ , such that the overall network has the form of a ring. Global best is selected among particle's neighbors. In another variant of PSO which is named fully informed PSO (FIPs) [20], each particle is not influenced by the best particle among its

neighbors. Instead, all its neighbors are used to update the velocity. After updating velocity and position of each particle, it is expected that swarm moves toward global solution. The generation updating continues to meet a stopping criterion such as predefined number of objective function evaluations.

### C. *k*-means clustering

Clustering is a task of dividing the data points into a number of clusters such that data points in the same clusters have high similarity while they are very dissimilar to entities in other clusters [21]. A clustering task needs two components to be determined, clustering algorithm and utilizing similarity measure. *k*-means clustering [22] is a well-known and simple clustering algorithm which puts data points into *k* groups based on a similarity measure. This method needs a predefined number of clusters to start clustering. At the first, *k* points are randomly selected among data points as initial centres of clusters. Then, a similarity measure like Euclidean distance is applied to assign each data point to corresponding nearest center. So *k* initial clusters have been constructed. Next step is the updating of centers and renewing the clusters. So that the mean of data points of each cluster are computed as new centers. In order to update clusters based on new centroids, the distance of each data point is computed from all centers and it is assigned to nearest cluster. These steps continues until the convergence criterion such as predefined iterations or no change met. The pseudo-code of the algorithm is presented in the Algorithm 1.

**Input :**  $D = \{x_1, x_2, x_3, \dots, x_n\}$ : Set of data points which should be clustered,  $K$ : Number of clusters  
**Output:**  $C = \{c^{(1)}, c^{(2)}, c^{(3)}, \dots, c^{(K)}\}$ : Set of clusters  
Initialize the centroids of clusters,  $\mu_1, \mu_2, \mu_3, \dots, \mu_K$  randomly  
**repeat**  
    Assign each data point to closest cluster based on Euclidean distance:  $c^{(i)} = \arg \min_j \|x^{(i)} - \mu_j\|^2$   
    Calculate new centroid for each cluster:  
     $\mu_j = (\sum_{c^{(i)}=j} x^{(i)})/n_j$ ; //  $n_j$  is the number of members in *j*th cluster  
**until** Convergence criteria met (no change in clusters);  
**Algorithm 1:** Pseudo-code for K-means algorithm

## III. PROPOSED METHOD

Collective intelligence of population in metaheuristic algorithms can yield to produce more elite solutions. In population-based algorithms, it is desirable that the population moves toward region of optimal solution but in most of them, each individual tries to enhance its own position. Accordingly, after producing better candidate solution, the previous one is replaced with new one. However, in some algorithms there is collaboration between a part of individuals to assist each other to produce better individual, utilizing collective intelligence of whole population is not proposed. As mentioned previously,

as a sample of collaboration between individuals, we can mention producing offspring using two or three individuals in most of evolutionary algorithms such as DE or Genetic Algorithms [23]. In PSO algorithm, collaboration can be defined as effect of global and local best solutions to update the position of each individual, but in all these cases, there is a partial collaboration of individuals. In this section, the details of the proposed scheme and its embedding in population-based algorithms are provided.

### A. Collective intelligence strategy

In this paper, collective intelligence of population is employed to increase elitism by producing generally a better candidate solution (CI-based trial vector). The method utilizes continues voting between individuals. The value of each variable of CI-based vector is produced using averaging between individuals that are more similar to each other in corresponding variable. Since in population-based algorithms, the individuals move toward better region of search space i.e, close to optimum point, so it is expected that among generations, the variables achieve better values to optimize objective function. Thus, voting among the variables of individuals creates a novel candidate solution which collects the intelligence of the whole population. In other words, it is preferred to construct individual with variable values which are obtained based on sort of aggregation established by all members of the population. The regular operators of algorithms (e.g., mutation and crossover in evolutionary algorithms) assemble good solutions in a slow manner (if they do) because a part of intelligence of population construct new individuals while by using generating CI-based vector (e.i., collecting valuable knowledge of all individuals at once), the process will be accelerated.

For this purpose, a new candidate solution is generated as follows. Suppose population includes the individuals,  $\langle x_1, x_2, x_3, \dots, x_N \rangle$  and each individual,  $x_i$ , is a *D* dimensional vector,  $x_{i,1}, x_{i,2}, x_{i,3}, \dots, x_{i,D}$ . For each dimension, *j*, *k*-means clustering is performed on individuals in the population based on only the value of corresponding variable. It means that a one dimensional *k*-means clustering is performed on  $x_{1,j}, x_{2,j}, x_{3,j}, \dots, x_{N,j}$  which are the one dimensional points related to *j*-th variable. So that for each variable *j*, the population is grouped into *K* clusters based on the similarity of individuals in corresponding variable. Therefore, individuals with similar values of variable are collected in the same cluster. The cluster with maximum number of members shows strong agreement on that variable. Accordingly, creating a new candidate based on the values of individuals in the most crowded cluster will help the search algorithm in exploration phase. The value of variable *j* in new candidate solution is acquired by averaging the values of variable *j* in members of most crowded cluster. It means that the continuous voting between individuals decides for the value of variables in new trial vector. Consequently, on each dimension, one-dimensional *k*-means clustering is applied to construct clusters of individuals and then selecting most crowded cluster determines the value of corresponding variable in CI-based individual. It is worth to

mention that since  $k$ -means clustering is done per dimension, however, the time complexity increases by increasing the dimension of the problem, it doesn't affect the  $k$ -means efficiency. In other words, this is a clustering for each dimension not a clustering method for whole population in the search space. For whole process, one-dimensional clustering would be applied. Fig. 1 illustrates the process of generating a trial vector using proposed collective intelligence-based method. In this example, population has 10 individuals and the most crowded cluster is indicated by gray color. In the case with same number of members in the clusters, a random cluster would be selected. As another perspective, the clustering process and generating CI-based vector in two dimensional search space is presented in Fig. 2. As can be seen, the value of variables on each dimension are clustered to three clusters, average of values in most crowded cluster would be the value of CI-based vector.

Similar to other population-based algorithms, optimization process begins with a randomly generated initial population. At each iteration, new individuals are generated by using operators or current individuals are updated. Based on the selection strategy of each algorithm, updated population as the next generation should move to the next iteration. According to proposed method, at the end of each iteration, a new trial vector is produced based on the collective intelligence scheme. The generated individual can be entered into population as a candidate solution if it is good enough. The trial vector is compared with the worst candidate solution in the population and it is replaced by the new vector if it has worse objective value. The detail of replacement strategy will be discussed in the following subsection.

In the optimization algorithms, as process progresses, the population is expected to move toward the more promising regions; therefore, by trusting on the only available knowledge distributed over the population, the variables of generated vector would gain values which gradually become closer to the optimal values. In other word, the density of individuals in each dimension shows the promisiness of each region in search space. It means that we expect that majority of voters can take the correct decision especially when the number of iterations increases (in fact that is a common assumption for any democratic system) . It is worth mentioning that the proposed method does not affect on any other part of optimization algorithm. Therefore, the proposed method can be applied in any population-based algorithm and any type of optimization problem. It means, the proposed CI-based scheme is population level approach not an operation level one.

### B. Embedding proposed scheme in a population-based optimization algorithm: DE and PSO as case studies

The CI-based candidate solution can influence movement of whole population toward better regions. In this paper, DE and PSO algorithms are modified using our proposed method. Since mutation operator in DE algorithm generates a new individual based on other candidate solutions in population, the CI-based individual can affect on whole population when

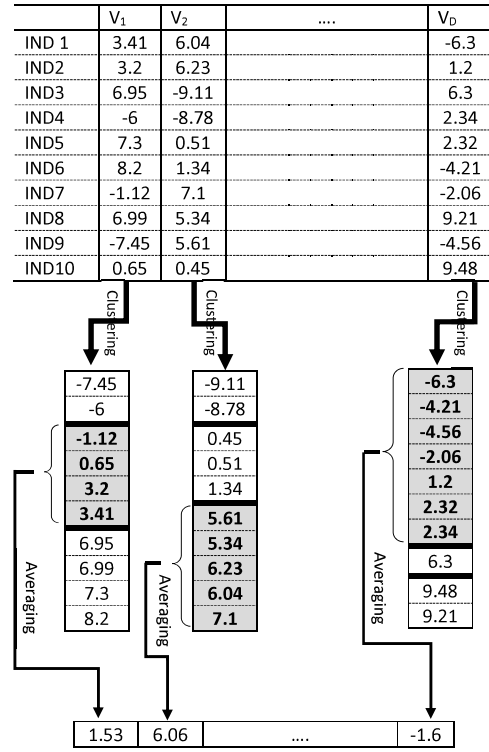


Fig. 1: An example of generating CI-based vector in a population with 10 individuals and  $D$  dimension.

it is selected as one of three randomly selected vectors for mutation operation. According to the proposed method, at each iteration of DE algorithm, the CI-based vector is constructed by collaborative intelligence of population and replace the worst candidate solution if the trial vector has a better objective value.

In the PSO algorithm, since each individual updates its own movement, the only effect which they get from others is related to updating the position based on the best candidate solution in their neighborhood. According to PSO updating equations, each individual determines its next position and velocity based on the best candidate solution in its neighborhood, so if it is desired to have more influence of the new CI-based vector on others, it is required to use a more contributed topology scheme to define neighborhood concept (like ring topology). For applying our method on PSO algorithm, at the end of each iteration, a new CI-based solution is generated based on the proposed scheme. Then, the new individual will be compared with the particle with worst personal best movement ( $x_w$ ). If the CI-based vector has better objective value comparing to worst personal best movement, the trail vector enters to the population and replaces  $x_w$ . The components of new particle (position and velocity) should be updated accordingly. CI-based particle gets the position based on voted variables (according to proposed scheme) and its personal best movement is initialized with its current position. But its velocity should

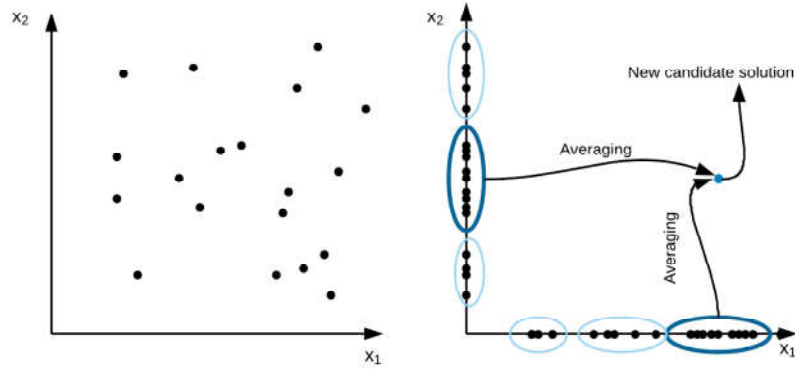


Fig. 2: Illustrating example of generating of CI-based vector in 2-dimensional space based on proposed scheme

be calculated at the next iteration according to the Eq. 2. However as a new member, its velocity be affected by only the global best. All steps of the proposed algorithm are provided in Algorithm 2.

#### IV. EXPERIMENTS RESULTS

In order to evaluate the performance of the proposed scheme, a series of experiments are conducted on CEC-2017 benchmark functions [24]. This test suite offers 30 scalable minimization problems, which can be categorized into: unimodal functions (F1, F2, F3), simple multi-modal functions (F4, F5, ..., F10), hybrid functions (F11, F12, ..., F20), and composition functions (F21, F22, ..., F30) [25]. The details of benchmark functions are provided in Table I.

As mentioned in the proposed method section, two meta-heuristic algorithms, DE and PSO are modified based on collective intelligence scheme, namely, CIDE and CIPSO. The proposed algorithms are compared with their parents (DE and PSO) on different dimensions which are valid for this benchmark set,  $D=10, 30, 50$ , and  $100$ . The same parameters are considered for all algorithms. The population size and number of function evaluations are set to  $100$  and  $30 \times D$ , respectively. The crossover rate,  $CR$ , and scaling factor,  $F$ , for DE and CIDV algorithms set to  $0.9$  and  $0.5$ . Optimal values of constriction coefficients and inertia weights for PSO proposed in [26] are  $0.7298$  and  $C_0=C_1=1.49618$ . The number of clusters ( $K$ ) in proposed scheme would be  $10$ . Each algorithm run  $30$  times. Then, the mean, median, best, worst and standard deviation of fitness function values for each algorithm are reported. To investigate the meaningful difference between algorithms, a Wilcoxon statistical test [27] with a confidence interval of  $95\%$  is performed on mean value of fitness functions. Symbols ' $\dagger$ ' and ' $\approx$ ' denote the CI version algorithms are better than and similar to compared algorithms, respectively. "w/t/l" indicates the number of "wins", "ties", and "loses" of proposed algorithm comparing to their parent version.

Tables II and III represent the resulted fitness values using DE algorithm and its CI-based version. Mean, median, best,

worst, and standard deviation of the results of  $30$  algorithm's runs are reported. All functions are solved in dimensions  $10, 30, 50$ , and  $100$ . The results confirm the significance improvement of proposed algorithm comparing to parent DE. The last row of the table indicates the result of statistical results. As it is presented, on dimension  $10$ , CIDE could outperform DE on  $18$  functions out of  $30$  while it loses on only  $3$  functions and it has same results with CIDE on  $8$  out of  $30$  functions. On other dimensions, the results are almost as dimension  $10$ . The difference among the results of dimension  $10$  and two higher dimensions ( $30$  and  $50$ ) is on the number of ties and loses. Regarding dimension  $100$ , the results of CIDE deteriorates slightly rather than lower dimensions while it is still has significant better results comparing to DE. According to obtained results, it is clear that CIDE performed almost same as DE for the unimodal problems (F1-F3) while for the multi-modal problems (F4-F10), it was able to obtain a solution with less distance to optimal solution for most of the functions (F5, F7, F8, F10) regard to the mean solutions comparing to DE. It means that the effect of injecting CI-based vector to population in such problems is more than unimodal ones. In the case of hybrid and composite functions, the CIDE obtained the closer solution to the optimal for some functions (F11, F17, F23, F24, F29) on all experimental dimensions. However, for other functions, CIDE performs differently on various dimensions.

The computational results for PSO and CIPSO are demonstrated in Tables IV and V. Similar to CIDE, CIPSO also outperforms its parent algorithm, PSO on most of the experimental functions. However, comparing to CIDE, CIPSO wins on less functions, it is still better on most of the functions compared to PSO. The important point about this comparison is that the improvement on results of CIPSO increases when dimension increases. So that on dimension  $10$ , CIPSO has better results on  $12$  functions out of  $30$ , while on dimension  $18$ , the number of wins is  $18$ . CIPSO ties with PSO on highest experimental dimension,  $100$ , on  $8$  functions out of  $30$  while the number of loses is only  $3$ . Regarding to type of functions,

```

input :  $NP$ : Population size,  $IT_{max}$ : Maximum
number of iterations,  $D$ : Number of
dimensions of the problem  $K$ : Number of
clusters
output: Best solution
// Initialization
Generate  $NP$  uniform random individuals to construct
 $POP_0$ ;
 $t = 1$ ;
Evaluate each individual in terms of fitness function;
while  $t < IT_{max}$  do
    // Run optimization algorithm on
     $POP_{t-1}$ 
     $POP_t \leftarrow \text{Optimize}(POP_{t-1})$ ;
    for  $d \leftarrow 1$  to  $D$  do // For each dimension
        // Performing the K-means
        clustering
         $C = \text{K-means}(POP_t(:, d)); // C =$ 
         $\{c^{(1)}, c^{(2)}, c^{(3)}, \dots, c^{(K)}\}$ 
        // Calculating the size of each
        cluster
         $NC = \{|c^{(1)}|, |c^{(2)}|, |c^{(3)}|, \dots, |c^{(K)}|\}$ ;
         $MC = \max(NC)$ ;
        // Averaging on variables in
        most crowded cluster
         $CIV(d) = \text{Average}(POP_t(c^{MC}, d))$ ;
    end
     $F_p = \text{Evaluate}(CIV)$ ;
    if  $DE$  is applied then
        Select candidate solution with worst fitness
        value ( $X_w$ );
        if  $F_p < \text{Evaluate}(X_w)$  then
             $POP_t(w) = CIV$ 
        end
    end
    if  $PSO$  is applied then
        Select candidate solution with worst personal
        best fitness value ( $X_{pw}$ );
        if  $F_p < \text{Evaluate}(X_{pw})$  then
             $POP(pw).position = CIV$ ;
             $POP(pw).pbest = CIV$ ;
        end
    end
     $t=t+1$ ;
end

```

**Algorithm 2:** Pseudo-code of proposed method; embedding proposed CI-based scheme in DE and PSO algorithms

TABLE I: Summary of the CEC 2017 test functions [24].  $N$  is the number of basic functions to construct the components of hybrid or composition functions.

	No.	Functions	$F_i(x^*)$
Unimodal	1	Shifted and Rotated Bent Cigar Function	100
	2	Shifted and Rotated Sum of Different Power Function*	200
	3	Shifted and Rotated Zakharov Function	300
Simple	4	Shifted and Rotated Rosenbrock's Function	400
	5	Shifted and Rotated Rastrigin's Function	500
Multimodal	6	Shifted and Rotated Expanded Scaffer's F6 Function	600
	7	Shifted and Rotated Lunacek Bi_Rastrigin Function	700
	8	Shifted and Rotated Non-Continuous Rastrigin's Function	800
Functions	9	Shifted and Rotated Levy Function	900
	10	Shifted and Rotated Schwefel's Function	1000
Hybrid	11	Hybrid Function 1 (N=3)	1100
	12	Hybrid Function 2 (N=3)	1200
	13	Hybrid Function 3 (N=3)	1300
	14	Hybrid Function 4 (N=4)	1400
	15	Hybrid Function 5 (N=4)	1500
	16	Hybrid Function 6 (N=4)	1600
	17	Hybrid Function 6 (N=5)	1700
	18	Hybrid Function 6 (N=5)	1800
	19	Hybrid Function 6 (N=5)	1900
	20	Hybrid Function 6 (N=6)	2000
Composition	21	Composition Function 1 (N=3)	2100
	22	Composition Function 2 (N=3)	2200
	23	Composition Function 3 (N=4)	2300
	24	Composition Function 4 (N=4)	2400
	25	Composition Function 5 (N=5)	2500
	26	Composition Function 6 (N=5)	2600
	27	Composition Function 7 (N=6)	2700
	28	Composition Function 8 (N=6)	2800
	29	Composition Function 9 (N=3)	2900
	30	Composition Function 10 (N=3)	3000
Search Range: $[-100, 100]^D$			

CIPSO is able to reach near to optimal solutions in unimodal and multi-modal functions on all dimensions. However, in hybrid and compositions functions, the difference of CIPSO performance is more explicit on higher dimensions.

For better illustration of performance evaluation, the convergence plots of proposed methods and parent algorithms are displayed in Fig. 3 and Fig. 4 on some selected functions. As it is presented, after the first few iterations, the CIDE and CIPSO algorithms surpass DE and PSO algorithms in terms of minimizing the fitness function. The injection CI-based vector to population improves the performance of the algorithm; so that the difference on some functions such as F12 to decreasing the error is very remarkable.

Table VI shows some impressive results to clarify how collective intelligence can generate a beneficial candidate solution. The table gives a comparison on quality of CI-based individual with best and worst individuals in the population. Each number indicates the percentage of the iterations that generated individual would be better than best or worst candidate solution in the population. The other column in this table is related to average rank of created individual in whole population. Comparing the quality of generated solution with best solution indicates that in some iterations of the algorithm, the new member will be injected to the population as the best solution (e.g., 8.3% in CIPSO on dimension 100). As it is presented, by increasing the dimension of the problem, the goodness of created individual comparing to worst solution increases. For example, in dimension 100, the result of voting is better than worst individual in 68% of iterations. These







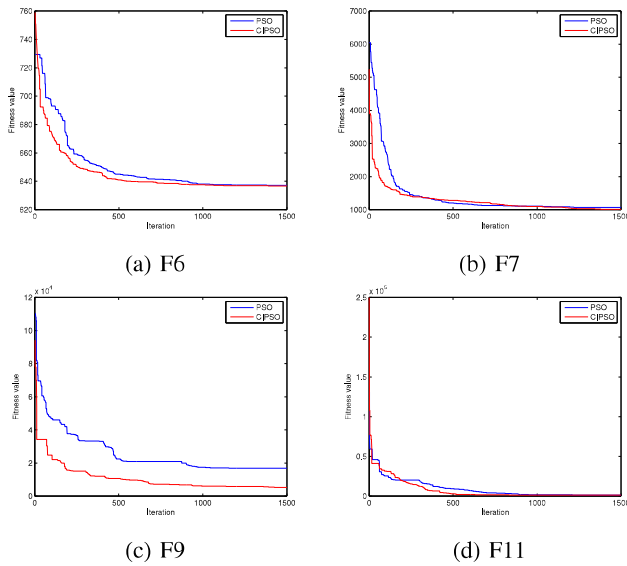


Fig. 4: Convergence plot for F6, F7, F9, F11 on D=50 for PSO and CIPSO.

TABLE VI: The quality of resulted candidate solutions using collective intelligence compared with best, worst, and whole population.

D	CIDE			CIPSO		
	Better than Best	Beter than Worst	AvgRank	Beter than Best	Beter than Worst	AvgRank
10	4.47	58.01	36.29	4.13	68.45	43.45
30	1.49	58.31	39.04	2.05	63.39	45.23
50	0.8945	54.43	42.12	1.4	70.02	60.82
100	7.4	68.51	53.67	8.3	79.12	65.34

tion of the new vector to population have a significant effect on improvement of acceleration rate of algorithms. In the present study, collective intelligence are applied in single-objective optimization while multi-objective case can be considered in future work. Moreover, as a future work, the effect level of intelligence for each individual can be defined according to its fitness value (e.g., utilizing weighted voting in generating new candidate solution).

## REFERENCES

- [1] K. Hussain, M. N. M. Salleh, S. Cheng, and Y. Shi, "Metaheuristic research: a comprehensive survey," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2191–2233, 2019.
- [2] S. Mahdavi, M. E. Shiri, and S. Rahnamayan, "Metaheuristics in large-scale global continues optimization: A survey," *Information Sciences*, vol. 295, pp. 407–428, 2015.
- [3] A. A. Freitas, "A survey of evolutionary algorithms for data mining and knowledge discovery," in *Advances in evolutionary computing*. Springer, 2003, pp. 819–845.
- [4] A. Chakraborty and A. K. Kar, "Swarm intelligence: A review of algorithms," in *Nature-Inspired Computing and Optimization*. Springer, 2017, pp. 475–494.
- [5] Y. Wang, "A sociopsychological perspective on collective intelligence in metaheuristic computing," *International Journal of Applied Metaheuristic Computing (IJAMC)*, vol. 1, no. 1, pp. 110–128, 2010.
- [6] R. Kicinger, T. Arciszewski, and K. De Jong, "Evolutionary computation and structural design: A survey of the state-of-the-art," *Computers & structures*, vol. 83, no. 23-24, pp. 1943–1978, 2005.
- [7] D. Whitley, "A genetic algorithm tutorial," *Statistics and computing*, vol. 4, no. 2, pp. 65–85, 1994.

- [8] S. P. T. P. Phyu and G. Srijuntongsiri, "Effect of the number of parents on the performance of multi-parent genetic algorithm," in *2016 11th International Conference on Knowledge, Information and Creativity Support Systems (KICSS)*. IEEE, 2016, pp. 1–6.
- [9] E. Bonilla-Huerta, J. C. H. Hernández, and R. Morales-Caporal, "A specialized random multi-parent crossover operator embedded into a genetic algorithm for gene selection and classification problems," in *Electrical Engineering and Control*. Springer, 2011, pp. 559–566.
- [10] L. M. Zheng, S. X. Zhang, K. S. Tang, and S. Y. Zheng, "Differential evolution powered by collective information," *Information Sciences*, vol. 399, pp. 13–29, 2017.
- [11] M. Elbes, S. Alzubi, T. Kanan, A. Al-Fuqaha, and B. Hawashin, "A survey on particle swarm optimization with emphasis on engineering and network applications," *Evolutionary Intelligence*, pp. 1–17, 2019.
- [12] O. Moslah, Y. Hachaichi, and Y. Lahbib, "Democratic inspired particle swarm optimization for multi-robot exploration task," 2016.
- [13] N. Monmarché, F. Guinand, and P. Siarry, *Artificial ants*. Wiley-ISTE Hoboken, 2010.
- [14] M. E. Aydin, J. Wu, and L. Zhang, "Swarms of metaheuristic agents: A model for collective intelligence," in *2010 International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. IEEE, 2010, pp. 296–301.
- [15] R. Eberhart and J. Kennedy, "Particle swarm optimization," in *Proceedings of the IEEE international conference on neural networks*, vol. 4. Citeseer, 1995, pp. 1942–1948.
- [16] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [17] S. Mahdavi, S. Rahnamayan, and A. Mahdavi, "Majority voting for discrete population-based optimization algorithms," *Soft Computing*, vol. 23, no. 1, pp. 1–18, 2019.
- [18] T. Tsujimoto, T. Shindo, T. Kimura, and K. Jin'no, "A relationship between network topology and search performance of pso," in *2012 IEEE Congress on Evolutionary Computation*. IEEE, 2012, pp. 1–6.
- [19] Y.-X. Wang and Q.-L. Xiang, "Particle swarms with dynamic ring topology," in *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 419–423.
- [20] C. W. Cleghorn and A. Engelbrecht, "Fully informed particle swarm optimizer: Convergence analysis," in *2015 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2015, pp. 164–170.
- [21] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Annals of Data Science*, vol. 2, no. 2, pp. 165–193, 2015.
- [22] V. Faber, "Clustering and the continuous k-means algorithm," *Los Alamos Science*, vol. 22, no. 138144.21, p. 67, 1994.
- [23] M. Srinivas and L. M. Patnaik, "Genetic algorithms: A survey," *computer*, vol. 27, no. 6, pp. 17–26, 1994.
- [24] N. Awad, M. Ali, J. Liang, B. Qu, and P. Suganthan, "Problem definitions and evaluation criteria for the cec 2017 special session and competition on single objective bound constrained real-parameter numerical optimization," *Technical Report*, 2016.
- [25] V. Kreischer, T. T. Magalhaes, H. J. Barbosa, and E. Krempser, "Evaluation of bound constraints handling methods in differential evolution using the cec2017 benchmark," in *XIII Brazilian Congress on Computational Intelligence, Rio de Janeiro, Brazil*, 2017.
- [26] J. Kennedy, R. Poli, and T. Blackwell, "Particle swarm optimization: an overview," *Swarm Intelligence*, vol. 1, no. 1, pp. 33–57, 2007.
- [27] F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in statistics*. Springer, 1992, pp. 196–202.