A Systematic Parameter Adaption Scheme in APSO

R.Sabin Begum[#1], Dr.R.Sugumar[*2]

[1]Research and Development center, Bharathiyar University, Coimbatore.
[2]Associate Professor, Department of Computer Science &Engg, Velammal Institute of Technology, Chennai.

***Abstract*—** An adaption feature of particle swarm optimization features have better search efficiency than particle swarm optimization (PSO) is presented. It can perform a global search over the entire search space with faster convergence speed. APSO enables automatic control of weight, acceleration coefficients, and other parameters to improve efficiency and convergence speed. . Results show that APSO substantially enhances the performance of the PSO paradigm in terms of convergence speed, and solution accuracy, and algorithm reliability.

***Keywords*—** Adaptive Particle Swarm Optimization; Global Optimization; Particle Swarm Optimization (PSO).

1. Introduction

Particle swarm optimization (PSO) was introduced by Kennedy and Eberhart in 1995 [1], [2], [3]. ThePSO have a simple mechanism that mimics swarm behavior inbirds flocking and fish schooling to guide the particles to searchfor globally optimal solutions.

Accelerating convergence speed and avoiding the local optimal have become the two most important and goals in PSO research.Adaptive PSO (APSO) is formulated by developing a systematic parameter adaptation scheme and an elitist learning strategy (ELS). To enable adaptation, an evolutionary state estimation (ESE) technique is first devised. Adaptive parameter control strategies can be developed based on the identified evolutionary state and by making use of existing research results on inertia weight [13][14] and acceleration coefficients.

2. PSO and its Developments

PSO Framework: In PSO, a swarm of particles are represented as potential solutions, and each particleI is associated with two vectors,i.e., the velocity vector

Vi=[v1iv2i,...,vDi]and the position vector Xi= [x1i,x2i,...,xDi], whereD stands for thedimensions of the solution space. The velocity and the positionof each particle are initialized by random vectors within thecorresponding ranges. During the evolutionary process, thevelocity and position of particle I on dimension d are updated as,

$$v_i^d = \omega v_i^d + c_1 rand_1^d (pBest_i^d - x_i^d) + c_2 rand_2^d (nBest^d - x_i^d) \quad (1)$$

$$x_i^d = x_i^d + v_i^d \quad (2)$$

Where ω is the inertia weight [13],c1 and c2 are the acceleration coefficients [2], andRand d1 and Rand d2 are two uniformly distributed random numbers independently generated within [0, 1] for the dth dimension [1]. In (1),pBestiis the position with the best fitness found so far for theith particle, and nBest is the best position in the neighborhood. In the literature, instead of using nBest,gBest may be used in the global-version PSO, whereaslBest may be used in the local-version PSO (LPSO).

3. ESE for PSO

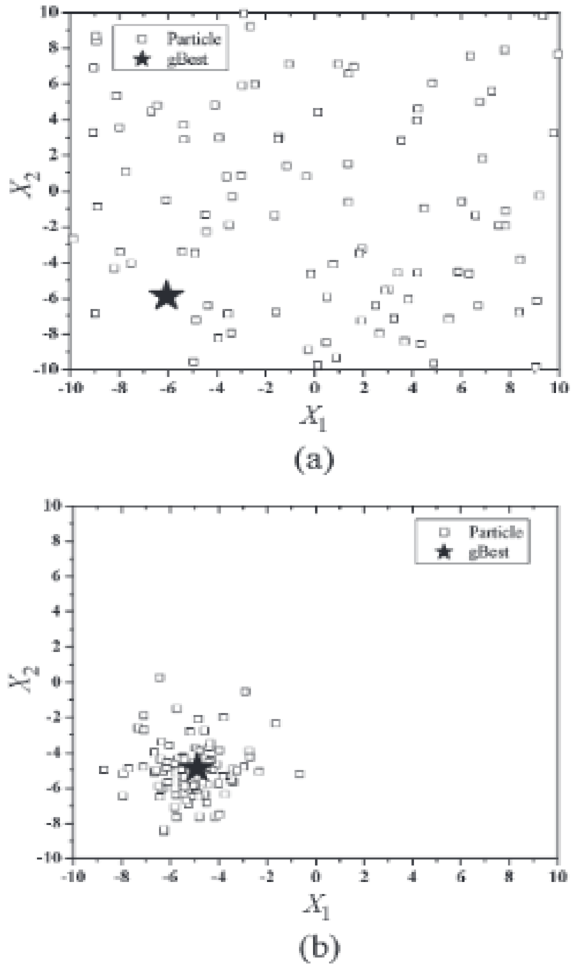

Fig.1: Population distribution observed at various stages in a PSO process. (a) Generation=1, (b) Generation=25

ESE: Based on the search behaviors and the population distribution characteristics of the PSO, an ESE approach is developed inthis section.

The sparsely distribution information in Fig. 1 can be formulated as illustrated in Fig. 2 by calculating the mean distance of each particle to all the other particles. It is reasonable to expect that the mean instance from the globally best particle to other particles would be minimal in the convergence state since the global best tends to be surrounded by the swarm. Incontrast, this mean distance would be maximal in the jumping out state, because the global best is likely to be away fromthe crowding swarm. Therefore, the ESE approach will take into account the population distribution information in every generation, as detailed given in the following steps.

Step 1: At the current position, calculate the mean distanceof each particle i to all the other particles.

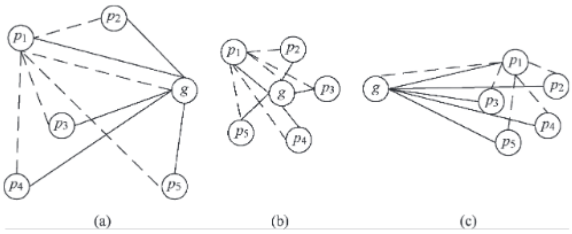

Fig.2: PSO population distribution information quantified by an evolutionary factor

For example, this mean distance can be measured using an Euclidian metric.

$$d_i = \frac{1}{N-1}\sum_{j=1, j\neq i}^{N}\sqrt{\sum_{k=1}^{D}\left(x_i^k - x_j^k\right)^2} \tag{3}$$

Where N and Dare the population size and thenumber of dimensions, respectively.

Step 2: Denotedof the globally best particle asdg .Compare alld I ’s, and determine the maximum andminimum distancesdmaxanddmin. Compute an“evolutionary factor”fas defined by

$$f = \frac{d_g - d_{\min}}{d_{\max} - d_{\min}} \in [0, 1]. \tag{4}$$

Take the time-varyingf1minimization processshown in Fig. 1 as an example to illustrate the variations off.

Step 3: Classifyfinto one of the four setsS1,S2,S3, andS4, which represent the states of exploration, exploitation, convergence, and jumping out, respectively.

4. APSO

4.1 Adaptation of the Inertia Weight

The inertia weightωinPSO is used to balance the global and local search capabilities.Many researchers have advocated that the value ofωshouldbe large in the exploration state and small in the exploitationstate [4], [13], [14]. However, it is not necessarily toDecreaseωpurely with time [14]. The evolutionary factorfshares some characteristics with the inertia weightω in thatfis also relatively large during the exploration state and becomesrelatively small in the convergence state.In this paper,ωis initialized to 0.9. Asωis not necessarilymonotonic with time, but monotonic withf,ωwill, thus, adaptto the search environment characterized byf.

In a jumping-outor exploration state[12], the largefandωwill benefit the globalsearch, as referenced earlier. Conversely, whenfis small, anexploitation or convergence state is detected, and, hence,ωdecreases to benefit the local search.

4.2 Control of the Acceleration Coefficients

Adaptive control can be devised for the acceleration coefficients based onthefollowing notion. Parameterc1represents the “self-cognition”that pulls the particle to its own historical best position, helpingexplor and maintaining the diversity of the swarm. Parameterc2 represents the “social influence” that pushes theswarm to converge to the current best regionglobally, helpingwith fast convergence [4], [12]. These are two differentmechanisms should be given different treatments in different evolutionary states [10]. In this work, the accelerationcoefficients areinitialized to 2.0 and adaptively controlledaccording to the evolutionary state, with strategies developed asfollows.

Strategy 1-Increasingc1and Decreasingc2 in an Exploration State: It is important to explore as many optima aspossible in the exploration state. Hence, increasingc1 and decreasingc2can help particles explore individually and achievetheir own historical best positions, rather than crowd around thecurrent best particle that is likely to be associated with a local
optimum.

Strategy 2-Increasingc1slightly and Decreasing c2Slightly in an Exploitation State: In this state, the particles are makinguse of local information and grouping toward possible localoptimal niches indicated by the historical best position ofeach particle. Hence, increasingc1slowly and maintaining arelatively large value can emphasize the search and exploitationaroundpBesti[9]. In the meantime, the globally best particle doesnot always locate the global optimal region at this stage yet.Therefore, decreasingc2slowly and maintaining a small valuecan avoid the deception of a local optimum. Further, an exploitation state is more likely to occur after an exploration stateand before a convergence state. Hence, changing directions forc1andc2should slightly be altered from the exploration stateto the convergence state.

Strategy 3-Increasingc1Slightly and Increasingc2Slightlyin a Convergence State: In the convergence state, the swarmseems to find the globally optimal region, and, hence, theinfluence ofc2should be emphasized to lead other particlesto the probable globally optimal region. Thus, the value ofc2should be increased. On the other hand, the value ofc1shouldbe decreased to let the swarm converge fast. However, extensiveexperiments on optimizing the 12 benchmark functions given in Table I revealed that such a strategy would prematurelysaturate the two parameters to their lower and upper bounds,respectively. The consequence is that the swarm will stronglybe attracted by the current best region, causing premature convergence, which is harmful if the current best region is a localoptimum. To avoid this, bothc1andc2are slightly increased.Note that, slightly increasing both acceleration parameterswill eventually have the same desired effect as reducingc1andincreasingc2, because their values will be drawn to around 2.0due to an upper bound of 4.0 for the sum ofc1andc2.

Strategy 4-Decreasingc1and Increasing c2in a Jumping-Out State: When the globally best particle is jumping out oflocal optimum toward a better optimum, it is likely to be faraway from the crowding cluster. As soon as this new region isfound by a particle, which becomes the (possibly new) leader,others should follow it and fly to this new region as fast aspossible. A largec2together with a relatively smallc1helpsto obtain this goal. Variations of the acceleration coefficients with the evolutionary state are illustrated in Fig. 3

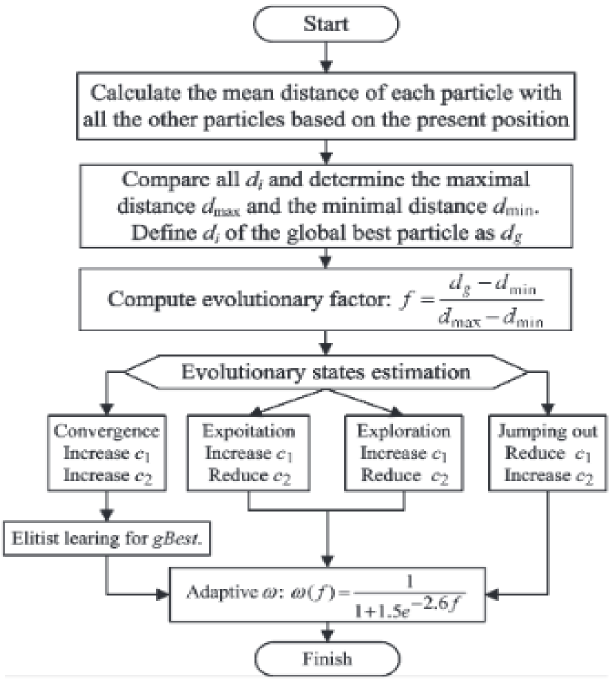

Fig. 3: ELS

The failures of using parameter adaptation alone for GPSOand VPSO on Schwefel's function suggest that a jumping-outmechanism would be necessary for enhancing the globality ofthese search algorithms. Hence, an "ELS" is designed here andapplied to the globally best particle so as to help jump out oflocal optimal regions when the search is identified to be in aconvergence state[5].Unlike the other particles, the global leader has no exemplarsto follow. It needs fresh momentum to improve itself. Hence, aperturbation-based ELS is developed to helpgBestpush itselfout to a potentially better region. If another better region isfound, then the rest of the swarm will follow the leader to jumpout and converge to the new region.The ELS randomly chooses one dimension ofgBest's historical best position, which is denoted byPdfor thedth dimension. Only one dimension is chosen because the local optimumis likely to have some good structure of the global optimum,and, hence, this should be protected. As every dimension hasthe same probability to be chosen, the ELS operation canbe regarded to perform on every dimension in a statisticalsense. Similar to simulated annealing, the mutation operationin evolutionary programming or in evolution strategies, elitistlearning is performed through a Gaussian perturbation.

$$P^d = P^d + (X^d_{\max} - X^d_{\min}) \cdot Gaussian(\mu, \sigma^2). \quad (5)$$

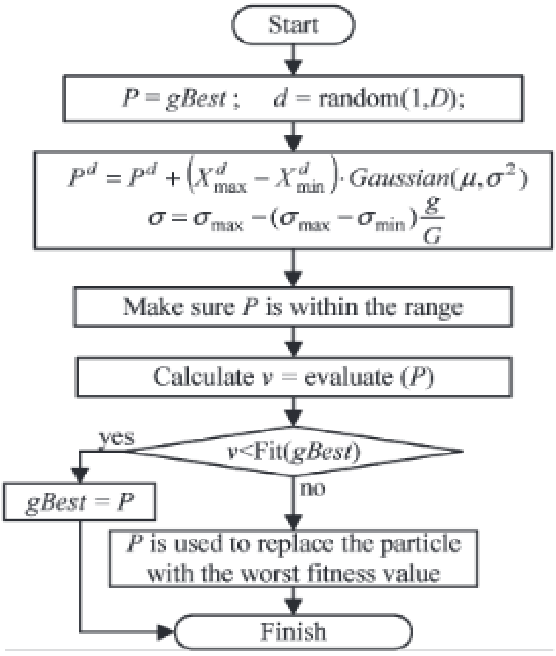

Fig.4: Search Behaviors of APSO

The search range[Xdmin,Xdmax]is the same as the lower and upper bounds of theproblem. TheGaussian(μ, σ2) is a random number of a Gaussian distribution with a zero meanμand a standard deviation (SD)σ, which is termed as the "Elitistlearning rate." Similar to some time-varying neural networktraining schemes, it is suggested thatσbe decreasedwith the generation number, which is given by[8]

$$\sigma = \sigma_{\max} - (\sigma_{\max} - \sigma_{\min})\frac{g}{G} \tag{6}$$

Where σmax and σmin are the upper and lower bounds ofσ,which represents the learning scale to reach a new region. Empirical study shows thatσmax=1.0andσmin=0.1 result in good performance on most of the test functions (referto Section VI-C for an in-depth discussion). Alternatively,σmay geometrically be decreased, similar to the temperature-decreasing scheme in Boltzmann learning seen in simulatedannealing. The ELS process is illustrated in Fig.4.

The complete flowchart of the APSO algorithm with adaptive parameters and ELS is shown in Fig.5.

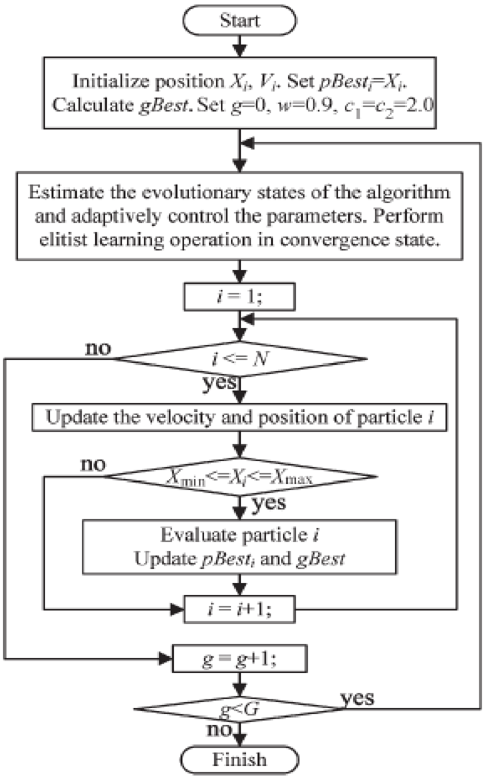

Fig.5: Adaptive parameters and ELS

Before applying theAPSO to comprehensive tests on benchmark functions, we firstinvestigate its search behaviors in unimodal and multimodal search spaces.

APSO in Unimodal Search Space: The search behaviorof the APSO in a unimodal space has been investigated onthe Sphere function (*f*1)[6]. In a unimodal space, it isimportant for an optimization or search algorithm to convergefast and to reine the solution for high accuracy. The inertiaweight confirms that the APSO maintainsa large ω in the exploration phase (for about 50 generations),and then a rapidly decreasing ω follows exploitation, leading toconvergence, as the unique global optimum region is found bya leading particle, and the swarm follows it. The ESE in APSO has influenced theacceleration coefficients. The curves for *c*1 and *c*2 somewhatshow good agreement with the ones given in Fig. 5. It can beseen that *c*1 increases while *c*2 decreases in the exploration andexploitation phases. Then, *c*1 and *c*2 reverse their directionswhen the swarm converges, eventually returning to around2.0. Then, trials in elitist learning perturb the particle thatleads the swarm, which is reflected in the slight divergencebetween *c*1 and *c*2 that follows. The search behavior on theunimodal function indicates that the proposed APSO algorithmhas indeed identified the evolutionary states and can adaptivelycontrol the parameters for improved performance.

APSO in Multimodal Search Space: Here, the APSO istested again to see how it will adapt itself to a multimodal space.When solving multimodal functions, a search algorithm shouldmaintain diversity of the population and search for as manyoptimal regions as possible. The search behavior of the APSO isinvestigated on Rastrigin's function (*f*8)[4],[7]. To comparethe diversity by the APSO and the traditional PSO,a yardstick proposed in [11] is used here, called the "populationstandard deviation," (psd)

$$psd = \sqrt{\left[\sum_{i=1}^{N}\sum_{j=1}^{D}\left(x_i^j - \overline{x}^j\right)^2\right] \Big/ (N-1)} \tag{7}$$

Where *N*, *D*, and *x* are the population size, the number of dimension,and the mean position of all the particles, respectively.

The variations in *psd*can indicate the diversity level of theswarm. If *psd*is small, then it indicates that the populationhasclosely converged to a particular region, and the diversityof thepopulation is low. A larger value of *psd*indicates thatthe population is of higher diversity. However,it does not mean that a larger *psd*is always better than asmaller one because an algorithm that cannot converge mayalso present a large *psd*. Hence, the *psd*needs to be consideredtogether with the solution that the algorithm arrives at.The results of *psd*comparisons are plotted andthose of the evolutionary processes . It can beseen that the APSO has an ability out from the localoptima, which is reflected by the regained diversity of thepopulation, as revealed in with a steady improvementin the solution, the inertia weight and the acceleration coefficientbehaviors of the APSO, respectively. These plots confirm that,in a multimodal space, the APSO can also find a potentialoptimal region (maybe a local optimum) fast in an early phaseand converge fast with a rapid decreasing diversity due to theadaptive parameter strategies. However, if the current optimalregion is local, then the swarm can separate and jump out.Hence, the APSO can appropriately increase the diversity of the population so as to explore for a better region owing tothe ELS in the convergence state. This behavior with adaptivepopulation diversity is valuable for a global search algorithm is used toprevent from being trapped in the local optima and to find the global optimum in a multimodal space.

5. Analysis of Parameter Adaptationand Elitist Learning

5.1 Merits of Parameter Adaptation and Elitist Learning

To quantify the significance of these two operations, the performance of APSO without parameter adaptation or elitistlearning is tested under the same running conditions . Results of the mean values on 30 independent trials are presented in[2],[10]. It is clear from the results that with elitist learning aloneand without adaptive control of parameters, the APSO can stilldeliver good solutions to multimodal functions. However, theAPSO suffers from lower accuracy in solutions to unimodalfunctions. As algorithms can easily locate the global optimalregion of a unimodal function and then refine the solution,the lower accuracy may be caused by the slower convergencespeed to reach the global optimal region. On the other hand,the APSO with parameter adaptation alone but without ELScan hardly jump out of the local optima and, hence, resultsin poor performance on multimodal functions. However, it canstill solve unimodal problems well.Note that both of the reduced APSO algorithms generallyoutperform a standard PSO that involves neither adaptationparameters nor elitist learning. However, the full APSO is themost powerful and robust for any tested problem. This is mostevident in the test results on $f4$. These results together with theresults confirm the hypothesis that parameteradaptation speeds up the convergence of the algorithm andelitist learning helps the swarm jump out of the local optimaand find better solutions.

5.2 Sensitivity of the Acceleration Rate

The effect of the acceleration rate, which is reflected byits bound δ, on the performance of the APSO is investigatedhere. For this, the learning rate σ is, hence, fixed (e.g., σmax =σmin = 0.5), and the other parameters of the APSO remain thesame as in Section V-A. The investigation consists of six teststrategies for δ, the first three being to fix its value to 0.01, 0.05,and 0.1, respectively, and the remaining three being randomlyto generate its value using a uniform distribution within [0.01,0.05], [0.05, 0.1], and [0.01, 0.1], respectively. The resultsare presented in Table X in terms of the mean values of thesolutions found in 30 independent trials.It can be seen that APSO is not very sensitive to the accelerationrate δ, and the six acceleration rates all offer goodperformance. This may be owing to the use of bounds for theacceleration coefficients and the saturation to restrict their sumby (12). Therefore, given the bounded values of $c1$ and $c2$ andtheir sum restricted by (12), an arbitrary value within the range[0.05, 0.1] for δ should be acceptable to the APSO algorithm.

6. Conclusion

In this paper, PSO has been extended to APSO. This progressin PSO has been made possible by ESE, which utilizes thepopulation distribution information and fitness of relative particle ,sharing a similar spirit to the internal modelling in evolutionstrategies. ESE-based parameter adaptation technique departs fromthe existing parameter variation schemes ,based on the generation number alone. Hence, the APSO is still simple and almost as easy to use asthe standard PSO, whereas it brings in substantially improvedperformance in terms of convergence speed, global optimality,solution accuracy, and algorithm reliability.

References

[1] J. Kennedy and R. C. Eberhart, “Particle swarm optimization,” in Proc. IEEE Int. Conf. Neural Netw., Perth, Australia, 1995, vol. 4,pp. 1942–1948.

[2] R. C. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in Proc. 6th Int. Symp. Micromachine Human Sci., Nagoya,Japan, 1995, pp. 39–43.

[3] J. Kennedy, R. C. Eberhart, and Y. H. Shi, Swarm Intelligence.San Mateo, CA: Morgan Kaufmann, 2001.

[4] R. C. Eberhart and Y. H. Shi, “Particle swarm optimization: Developments, applications and resources,” in Proc. IEEE Congr. Evol.Comput.,Seoul, Korea, 2001, pp. 81–86.

[5] X. D. Li and A. P. Engelbrecht, “Particle swarm optimization: An introduction and its recent developments,” in Proc. Genetic Evol. Comput.Conf., 2007, pp. 3391–3414

.[6] R. A. Krohling and L. dos Santos Coelho, “Coevolutionary particle swarmoptimization using Gaussian distribution for solving constrained optimization problems,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 36, no. 6, pp. 1407–1416, Dec. 2006.

[7] N. Franken and A. P. Engelbrecht, “Particle swarm optimization approachesto coevolve strategies for the iterated prisoner’s dilemma,” IEEE Trans. Evol.Comput., vol. 9, no. 6, pp. 562–579, Dec. 2005.

[8] S.-Y. Ho, H.-S. Lin, W.-H.Liauh, and S.-J. Ho, “OPSO: Orthogonal particle swarm optimization and its application to task assignment problems,” IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 38, no. 2, pp. 288–298, Mar. 2008.

[9] B. Liu, L. Wang, and Y. H. Jin, “An effective PSO-based memetic algorithmfor flow shop scheduling,” IEEE Trans. Syst., Man, Cybern. B,Cybern., vol. 37, no. 1, pp. 18–27, Feb. 2007.

[10] R. C. Eberhart and Y. Shi, “Guest editorial,” IEEE Trans. Evol. Comput.—Special Issue Particle Swarm Optimization, vol. 8, no. 3, pp. 201–203, Jun. 2004.

[11] G. Ciuprina, D. Ioan, and I. Munteanu, “Use of intelligent-particle swarmoptimization in electromagnetics,” IEEE Trans. Magn., vol. 38, no. 2, pp. 1037–1040, Mar. 2002.

[12] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, “Comprehensive learning particle swarm optimizer for global optimization of multimodal functions,” IEEE Trans. Evol. Comput., vol. 10, no. 3, pp. 281–295, Jun. 2006.

[13] Y. Shi and R. C. Eberhart, “A modified particle swarm optimizer,” in Proc.IEEE World Congr.Comput.Intell., 1998, pp. 69–73.

[14] Y. Shi and R. C. Eberhart, “Empirical study of particle swarm optimization,”Proc. IEEE Congr.Evol.Comput., 1999, pp. 1945–50.

[15] A. Chatterjee and P. Siarry, “Nonlinear inertia weight variation for dynamicadaptation in particle swarm optimization,” Comput.Oper. Res.,vol. 33, no. 3, pp. 859–871, Mar. 2004.