

A PARTICLE SWARM OPTIMIZATION BASED CHAOTIC K-MEANS EVOLUTIONARY APPROACH

Min-Yuan Cheng

National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C
myc@mail.ntust.edu.tw

Kuo-Yu Huang

National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C
d9405502@mail.ntust.edu.tw

Hung-Ming Chen

National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C
hungming@mail.ntust.edu.tw

Abstract

The proposed approach brings up a manner of cognitiveness that inherits a paradigm in particle swarm optimization to implement a chaotic mapping and enhanced by K-means clustering algorithm. In this work, named KCPSO, chaotic mapping with ergodicity, irregularity and the stochastic properties in PSO contributes to global search while K-means with clustering properties in PSO results in rapid convergence.

Numerical results indicated that PSO adopting chaotic features can more easily escape from local optima, and meanwhile PSO incorporating with K-means can evidently improve convergence speed. Unpredictability and grouping principle underlying chaotic mapping and K-means methods often imply diversity maintenance and convergence potential within the swarm which inevitably lead to a desirable optimal solution.

As proven in the experiments with multidimensional search space and compared with original PSO, the conclusions reported that the proposed KCPSO algorithm could improve the search performance on the benchmark functions significantly, and show the effectiveness of solving optimization problems.

KEYWORDS: Particle swarm optimization, Chaotic mapping, K-means

INTRODUCTION

Particle swarm optimization (PSO) was first introduced in 1995 (Kennedy and Eberhart, 1995). PSO algorithm has been developing very rapidly and many results have been reported. It is characterized as a simple concept, which is robust and easy to implement and compute. Concerning its implementation, PSO can easily be programmed and has few parameters to regulate. Due to these superior features of PSO algorithm, it has been emerged as a new and attractive optimization tool that has successfully applied in a variety of different fields.

Nowadays, PSO has been successfully applied to various optimization problems (Shi and Eberhart, 1999; Shi and Eberhart, 1998).

Some reported works focused on the modification of PSO to solve the parameter selection problems (Zhao and Yang, 2009; Lin S.W., et al., 2008; Wu et al., 2008; Shen et al., 2007; Yuan and Chu, 2007). In 2008, a PSO algorithm was proposed to solve the parameter estimation problems for non-linear dynamic rational filters (Lin Y.L., et al., 2008). The proposed approach had significantly improved the approximation results compared with GA. On the other side, PSO combined chaos to enhance the searching efficiency had greatly improved the evolutionary abilities (Liu et al., 2005). In recent years, chaotic sequences have been adopted instead of random sequences and very interesting and somewhat good results have been shown in many applications (Alatas et al., 2009) such as secure transmission (Suneel, 2006; Wong et al., 2005), and nonlinear circuits (Arena et al., 2000), DNA computing (Manganaro and Pine, 1997), image processing (Gao et al., 2006).

Clustering is one of the most important and the most challenging of classifying algorithms. A successful clustering algorithm is able to reliably find true natural groupings in the data set. K-means is one of the well-known algorithms for clustering, originally known as Forgy's method (Forgy, 1965). K-means clustering is the process of dispatching a set of objects into groups or clusters of similarities. Objects collected in the same cluster have similar features, but others are not (Han & Kamber, 2001). K-means is famous for its simplicity and computational efficiency in clustering techniques. As aforementioned Chaotic algorithm is for population diversity in PSO, and K-means is for convergence efficiency in evolution. The former will keep the system accuracy, and the later will decrease iteration times of PSO significantly.

In the face of increasing complexity and dimensionality of coming applications coupled with their tendency of premature convergence due to the high convergence speeds, some investigations have been undertaken to improve the performance of conventional PSO (Alatas and Akin, 2009; Leandro and Antonio, 2009; Leandro and Viviana, 2009). But, little attention is given in chaos coupled with K-means and PSO. And then, a hybrid PSO with adaptive chaotic search and K-means algorithm is proposed to optimize the object function in this paper. The proposed competitive K-means and chaos evolutionary approach is adapted for particle swarm optimization algorithm design, named KCPSO, which appears to have considerable potential for solving complex optimization problems.

The rest of the paper is organized as follows. In section 2, descriptions of K-means, chaos and PSO are presented. In Section 3, the proposed algorithm, KCPSO, is described and analyzed in detail. The computational results are presented and analyzed in Section 4. In the last section, conclusions and future research are given.

APPROACHES FOR OPTIMIZATION

K-means Clustering Approach

Clustering is the process of grouping a set of physical or abstract items into clusters by similar features. K-means is one of the well-known algorithms for clustering, and it has been employed extensively in various fields including exploring studies: such as data mining, statistical data analysis: such as Custom Relationship Management, and other business

applications. The K-means algorithm for clustering is based on the mean value of items in the group. It is suggested to assign each item to the cluster with the nearest centroid (mean) (Mac Queen, 1967). In general, in this study the primary operating procedures for K-means are presented as follows:

- (1) Defining how many clusters are to be created.
- (2) Randomly assigning initial items to different clusters.
- (3) Assigning new items to the cluster whose location to centroid is the nearest (by Euclidean distance with either standardized or un-standardized observations) and re-calculate the centroid for the existing or updated clusters.
- (4) Repeating Step (3) until no more reassigning.

Chaos Approach

Chaos can be considered traveling particles within a limited range occurred in a deterministic nonlinear dynamic system. There is no definite regularity for such a traveling path. Such a movement is very similar to a random process, but extremely sensitive to the initial condition. Chaotic dynamic mappings have been defined as noninvertible mappings of the (0, 1) interval onto itself. Logistic mapping (May, 1976; Feigenbaum, 1978) is one of the most important Chaotic dynamic mappings which defines the simplest mapping for studying the period-doubling bifurcation (vide infra). In the well-known logistic equation (May, 1976):

$$X_{n+1} = f(\mu, X_n) = \mu X_n (1 - X_n) \quad (1)$$

in which μ stands for a control parameter, X for a variable and $n = 0, 1, 2, 3, \dots$. It is easy to find that equation (1) is a deterministic dynamic system. The variable X is also called as chaotic variable. The basic characteristic of chaos could be presented by Eq. (1), for a very small difference in the initial value of X will cause large difference in its long-term behavior.

The variation of control parameter μ of Eq. (1) will directly impact the behavior of X greatly. Usually, $[0, 4]$ has been defined as domain area of control parameter μ . Different value in domain area of μ will determine whether X stabilizes at a constant size or behaves chaotically in an unpredictable pattern. The track of chaotic variable looks like in disorder. However, it can travel ergodically over the whole space of interest especially under the condition of $\mu = 4$. Then, a tiny difference in initial value of the chaotic variable would result in considerable differences of the values of chaotic variable later. Generally, there are three primary characteristics of the variation of the chaotic variable, i.e. ergodicity, irregularity and pseudo-randomness (Bountis, 1995; Li & Jiang, 1998; Ohya, 1998).

Logistic equation as shown in equation (1) can be distinguished by four intervals in accordance with the value of μ . First, when the value of μ is smaller than 1.0, the chaotic variable X_{n+1} converges to a stable point 0.0. Then, if the value of μ is between 1.0 and 3.0, no matter what initial value for X_0 between 0.0 and 1.0 was taken, X_{n+1} would converge to a certain value between 0.0 and 0.63665. And, the bifurcation occurs from $\mu \geq 3.0$. The system will enter the chaos domain, if μ reaches a critical point of 3.5699456.... Finally, when $\mu = 4.0$ the values of X_{n+1} would take any real numbers between 0.0 and 1.0 and no redundant value would present again while having turned up already. In this study, ' $\mu = 4.0$ ' was taken to have the advantages of diversity during evolution.

Particle Swarm Optimization

PSO has constructive cooperation by sharing information with other particles (Liu et al., 2005). During iterative process, the behavior of each particle is compromised among three possible affecting factors: its own previous velocity and cognitive component and social component. The relatively high value of cognitive component will result in an excessive wandering of particle through search space, whereas a high value of social component may lead to a premature. In fact, the fitness value of each particle is determined by some defined functions and the velocity of each particle is updated by keeping track of global best position (*gbest*) and particle best position (*pbest*) in every iteration. Here, *pbest* is the best design of each particle traversed so far while *gbest* is the global best design found up to the current iteration. Therefore, the proper control of these variants, cognitive component *pbest* and social component *gbest*, can improve the performance of PSO.

The original PSO algorithm was developed by Kennedy and Eberhart in 1995. After many numerical simulations, Shi and Eberhart suggested adding a weighting factor to prevent premature convergence of the original PSO algorithm (Shi and Eberhart, 1998). Therefore, let $p_i = [p_{i1}, p_{i2}, \dots, p_{iD}]$ and $v_i = [v_{i1}, v_{i2}, \dots, v_{iD}]$ be D -dimensional vectors which present the position and velocity of i th particle along the search space, respectively. Eq. (2) and Eq. (3) are the modified particle velocity and position updating equation.

$$v_i(t+1) = w \times v_i(t) + c_1 \times r_1 \times (pbest_i(t) - p_i(t)) + c_2 \times r_2 \times (gbest_i(t) - p_i(t)) \quad (2)$$

$$p_i(t+1) = p_i(t) + v_i(t+1) \quad (3)$$

Where c_1 and c_2 are two acceleration constants, stand for the cognitive and social parameters, respectively; r_1 and r_2 are two positive random numbers uniformly distributed in the range [0, 1]. According to Kennedy and Eberhart, these cognitive and social parameters, c_1 and c_2 , are set equal to 2 in order to make the average velocity change coefficient close to 1 (Kennedy and Eberhart, 1995). In Eq. (2) w is inertia weight in the range [0,1], and the subscript i indicates the i th design variable, and t is the iteration number. v and p denote the velocity and position, respectively.

PROPOSED APPROACH KCPSO

Inspired by the chaos spread-spectrum distribution and unpredictable irregular motions and the *gbest* fluctuating migration in PSO, this paper provides a novel combined optimization method, which introduces chaotic search into PSO based on K-means clustering algorithm, named KCPSO, so as to improve the searching performance for global optimum. It is a promising way to achieve optimal trade-off between exploration and exploitation. Moreover, it is an effective way in dealing with the updating of velocity and position in KCPSO. The use of chaotic procedure combined with K-means clustering algorithm in PSO can be more helpful to solve some multimodal functions and avoid getting stuck from local optima than that can be done in a conventional PSO.

K-means plays a critical role in convergence of PSO. Chaos algorithm can keep PSO population diversity and avoid from premature. To take advantages of the above described benefits in PSO, a novel algorithm combined K-means clustering and the Chaos algorithms

with parallel population-based PSO was proposed as a powerful hybrid algorithm called KCPSO (K-means and Chaos in PSO Algorithm). Initial swarm members of KCPSO should be generated by chaotic algorithm, and then chaos function would keep diversifying the swarm members in an ergodicity, irregularity and pseudo-randomness way. Then, K-means clustering in this study will help to group swarm members in several clusters as pre-defined. Thus, location information of each centroid of cluster would be treated as candidate swarm members for the coming evolution. A competing procedure was employed to eliminate lower fitness value members, and reserved the others to create formal swarm for KCPSO iteration.

Chaos procedure activated in PSO procedure will contribute to locate global optimum in early stage, while K-means clustering activated in PSO procedure will improve convergence performance in final stage. The proposed PSO based approach employed the complementary properties from chaos and K-means clustering algorithm will present their mutual talents in exploration and exploitation.

Chaos is nonlinear in nature (Lu, Z. et al., 2006), and can help global searching by its diversity characteristic; however, it will cause much more computation time before getting convergent. To perform effectively, chaos algorithm employed in this study was a probability based function; that is, chaos function would be triggered with high probability in the beginning and decreasingly by iteration times. At the meantime, chaos function contributed global searching to PSO evolution from initial, and reduced its affecting range by times to comply with PSO convergent procedure at end.

K-means is a grouping technique which explores existing searching space, points out the centers of population groups and offers a short cut to convergence during PSO evolution. Unfortunately, K-means efficient convergent feature could not improve the heterogeneity of swarm members to enlarge the possible searching space for global optimization in this study. Hybrid chaos and K-means with PSO could combine their characteristics and merits together; therefore, a PSO based global searching with effectively convergent could be reached. Apparently, it is a nonlinear approach to search in global space with less fit for linear problems.

In this work, the initialized particles were generated by chaos procedure to diversify their positions between the lower and upper bound of the domain values. Moreover, the parallel population-based evolutionary characteristics of PSO, ergodic and irregular properties of chaos and potential moving tendency resulted by K-means clustering algorithm were cohesively combined to show their superior performance in locating optimal solution. During evolution, KCPSO would generate a certain rule to direct swarm migration, then swarm members would be modified coherently along the evolution process which represented the designed association rules. In particular, K-means clustering algorithm and chaotic approach would be exploited in KCPSO to thoroughly explore the entire search space and to point out the most appropriate migration way for the designed algorithm. Simulation results and comparisons had clearly demonstrated the effectiveness and efficiency of the proposed KCPSO. And, the flow chart of proposed algorithm KCPSO is described as following:

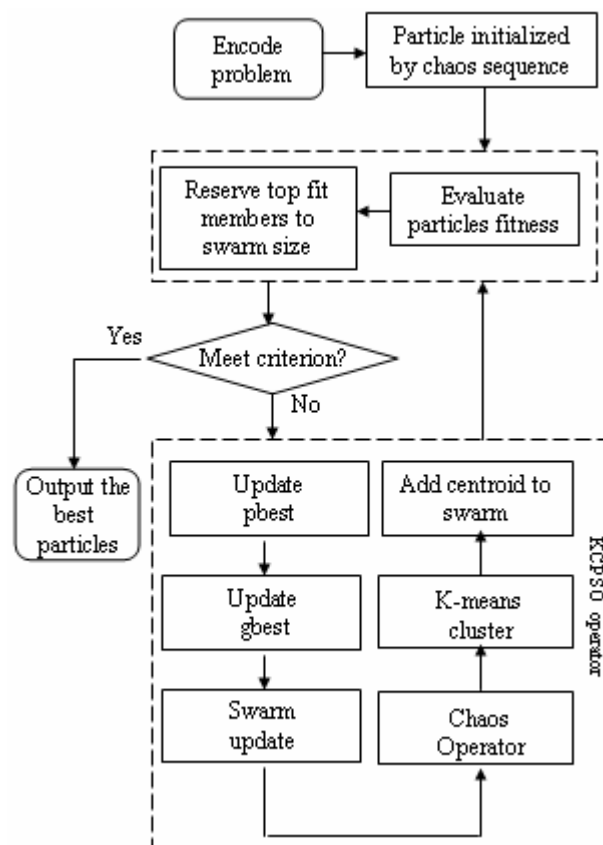


Fig. 4. Flowchart of KCPSO algorithm

EXPERIMENTS AND ANALYSIS

Experiment Setup

All the experiments were completed on a 1.83 GHz Core 2 Due T5550 processor with 2.99 GB of Random Access Memory (RAM). In every case study, 50 independent runs were made for each of the optimization methods involving 40 different initial particle trial solutions for every optimization method. For each testing problem, the parameters were set as follows: $c_1 = c_2 = 2.05$. Two criteria were applied to terminate the computation work: reaching maximum number of iterations '1E+5', and the second criterion was that swarm members converged within a minimum scope of 1E-05.

Experiments Over KCPSO

KCPSO was designed to benefit from the well constructed scheme of complementary contributions of the employed algorithms. Experiments have been performed to evaluate and compare the outcomes of KCPSO with other PSO methods. In Table 1-4, for every 50 independent runs, KCPSO takes the shortest average evaluation time and find out global optimum with excellent probability. Accordingly, it is concluded that KCPSO is much effective and robust for these kinds of complex numerical optimizations.

The proposed chaotic search coupled with PSO can not work well in cope with Sphere and Rosenbrock function, while K-means clustering algorithm embedded in PSO has poor performance in solving Rastrigrin and Griewank functions, as shown in Table 1-4. However, the complementary characteristics of chaos, K-means and PSO are well arranged in KCPSO, their contributions have improved the performance of PSO to superior which can be easily proved by an analysis of the computational results in this work.

In this context, the use of chaotic search in PSO was more helpful to conduct a proper way for swarm migration than that could be done through the K-means clustering algorithm in conventional PSO. Meanwhile, K-means clustering in PSO would conduct unpredictable moving tendency to help escape from local optima. Consequently, the experiment results have shown that the well hybridized KCPSO has earned an excellent synergy over combination in terms of fast convergence and perfect success rate. The outstanding performance also confirms the effectiveness and robustness of the proposed method.

Table 1-4 illustrate the varying performances resulted by different approach. It can be seen that the impact of iteration over success rate is more sensitive for KCPSO than for other three algorithms. The computational results have clearly demonstrated that KCPSO, employing complementary features from K-means clustering algorithm, chaos approach and PSO, is a well designed algorithm for solving unimodal and multimodal problems.

Table 1: Sphere Function

Functions	KCPSO		KPSO		CPSO		PSO	
	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.
Success		1.00		1.00		0.70		0
Mean	103.80	0.55	3034.75	25.80	35027.60	44.55	8332.20	18.89
Min	85	0.43	2332	20.85	1454	1.85	3705	8.53
Max	121	0.76	3623	30.63	100000	128.49	18669	41.87
Std. D.	10.29	0.08	347.58	2.78	44208.81	56.20	3888.20	8.65

Table 2: Rosenbrock Function

Functions	KCPSO		KPSO		CPSO		PSO	
	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.
Success rate		1.00		1.00		0		0
Mean	4845.65	26.81	4909.35	41.54	14891.55	20.31	100000	54.80
Min	4387	23.93	4064	35.46	6637	9.03	100000	54.72
Max	5489	30.00	5910	52.01	38440	52.45	100000	54.98
Std. D.	351.44	1.95	523.45	5.10	8046.33	10.97	0.00	0.06

Table 3: Rastrigrin Function

Functions	KCPSO		KPSO		CPSO		PSO	
	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.
Success rate		1.00		0		0.80		0
Mean	99.50	0.53	2600.20	22.18	26646.55	39.18	6072.90	5.32
Min	75	0.39	2133	17.62	1924	2.85	2778	2.43

Max	124	0.77	3407	28.65	100000	147.02	12168	10.71
Std. D.	11.95	0.09	309.29	2.85	38296.18	56.19	2620.13	2.30

Table 4: Griewank Function

Functions	KPSO		KPSO		CPSO		PSO	
	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.	Iteration	Sec.
Success	1.00		0.15		0.90		0	
Mean	105.80	0.65	2864.55	25.21	16103.95	35.24	4274.75	5.57
Min	82	0.51	2417	21.23	1178	2.62	2800	3.75
Max	129	0.89	3361	29.69	100000	218.88	6963	8.80
Std. D.	13.26	0.09	325.01	2.87	30052.47	65.55	1319.05	1.68

CONCLUSIONS

This study has proposed a procedure which joins K-means and chaos attributes based on particle swarm optimization algorithm. The proposed procedure is not only to enhance the diversity of PSO for more accuracy but also to extract clustering rules for achieving a potential trend of evolution. Additionally, it can effectively improve some drawbacks of traditional PSO, such as long running time and getting trapped in local optima.

However, the future works will include detailed theoretically investigations, such as parallel or distributed implementation of the designed algorithm and modifications for various heuristic methodologies (e.g., ant colony optimization, A.C.O.) to improve their performances.

REFERENCES

- Alatas Bilal, Akin Erhan. (2009) Chaotically encoded particle swarm optimization algorithm and its applications. *Chaos, Solitons & Fractals*, 41(2), 939-950.
- Alatas Bilal, Akin Erhan. Bedri Ozer. A., (2009) Chaos embedded particle swarm optimization algorithms. *Chaos, Solitons and Fractals*, 40, 1715–1734.
- Arena, P., Caponetto, R., Fortuna, L., Rizzo, A., La Rosa M., (2000) Self-organization in nonrecurrent complex systems. *International Journal of Bifurcation & Chaos in Applied Sciences & Engineering*, 10(5), 1115–1125.
- Bountis, T., (1995) Fundamental concepts of classical chaos. Part I. *Open Systems & Information Dynamics*, 3(1), 23-95.
- Eberhart, R.C., Shi, Y., (2001) Particle swarm optimization: developments, applications and resources. In: *Proceedings of IEEE international conference on evolutionary computation*, p. 81–86.
- Feigenbaum, M.J., (1978) Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, 19(1), 25–52.

- Forgy, E. (1965). Cluster analysis of multivariate data: Efficiency versus interpretability of classifications. *Biometrics*, 21, 768.
- Gao, H., Zhang, Y., Liang, S., Li, D.A., (2006) New chaotic algorithm for image encryption. *Chaos, Solitons & Fractals*, 29, 393–399.
- Han, J., Kamber, M., (2001) *Data mining: Concepts and techniques*. San Francisco: Morgan Kaufmann Publishers.
- Kennedy, J., Eberhart, R.C., (1995) Particle swarm optimization. In: *Proceedings of the IEEE international conference on neural networks*. Perth, Australia 4, 1942–1948.
- Leandro dos Santos Coelho, Antonio Augusto Rodrigues Coelho. (2009) Model-free adaptive control optimization using a chaotic particle swarm approach. *Chaos, Solitons & Fractals*, 41, 2001–2009.
- Leandro dos Santos Coelho, Viviana Cocco Mariani. (2009) A novel chaotic particle swarm optimization approach using Hénon map and implicit filtering local search for economic load dispatch. *Chaos, Solitons & Fractals*, 39, 510–518.
- Li, B., Jiang, W.S., (1998) Optimizing complex functions by chaos search. *Cybernetics and Systems*, 29(4), 409-419.
- Lin, S.W., Ying, K.C., Chen, S.C., Lee, Z.J., (2008) Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Systems with Applications*, 35(4), 1817–1824.
- Lin, Y.L., Chang, W.D., Hsieh, J.G., (2008) A particle swarm optimization approach to nonlinear rational filter modeling. *Expert System with Application*, 34, 1194–1199.
- Liu, B., Wang, L., Jin, Y.H., Tang, F., Huang, D.X., (2005) Improved particle swarm optimization combined with chaos. *Chaos, Solitons & Fractals*, 25(5), 1261–1271.
- Lu, Z., Shieh, L.S., Chen, G.R., Coleman, N.P., (2006) Adaptive feedback linearization control of chaotic systems via recurrent high-order neural networks. *Information Sciences*, 176(16), 2337-2354.
- Mac Queen, J.B., (1967) Some methods for classification and analysis of multivariate observations. In *Proceedings of 5th Berkeley symposium on mathematical statistics and probability*. Berkeley: University of California Press. p. 281–297
- Manganaro, G., Pine dade Gyvez J., (1997) DNA computing based on chaos. In: *Proceedings of the 1997 IEEE international conference on evolutionary computation*. Piscataway (NJ): IEEE Press; p. 255–260.
- May, R.M., (1976) Simple mathematical models with very complicated dynamics. *Nature*, 261, 459–467.
- Molga, M., Smutnicki, C., (2010), “Test functions for optimization needs”, available at: “<http://www.zsd.ict.pwr.wroc.pl/files/docs/functions.pdf>”

- Montiel, O., Castillo, O., Melin, P., (2007) Human evolutionary model: A new approach to optimization. *Information Sciences*, 177(10), 2075–2098.
- Mooney, A., Keating, J.G., Heffernan, D.M., (2006) A detailed study of the generation of optically detectable watermarks using the logistic map. *Chaos, Solitons & Fractals*, 30, 1088–1097.
- Niknam, T., (2006) An approach based on particle swarm optimization for optimal operation of distribution network considering distributed generators. In: *Proceedings of the 32nd Annual Conference on IEEE Industrial Electronics, IECON 2006*. p. 633–637.
- Niknam, T., (2010) A new fuzzy adaptive hybrid particle swarm optimization algorithm for non-linear, non-smooth and non-convex economic dispatch problem. *Applied Energy*, 87(1), 327-339.
- Niknam, T., Amiri, B., Olamaie, J., Arefi, A., (2009) An efficient hybrid evolutionary optimization algorithm based on PSO and SA for clustering. *Journal of Zhejiang University: Science A*, 10(4), 512-519.
- Ohya, M., (1998). Complexities and their applications to characterization of chaos. *International Journal of Theoretical Physics*, 37(1), 495-505.
- Shen, Q., Shi, W.M., Kong, W., Ye, B.X., (2007) A combination of modified particle swarm optimization algorithm and support vector machine for gene selection and tumor classification. *Talanta*, 71(4), 1679–1683.
- Shi, Y., Eberhart, R.C., (1998) A modified particle swarm optimizer. In: *Proceedings of the IEEE international conference on evolutionary computation*. Anchorage, Alaska. p. 69–73.
- Shi, Y., Eberhart, R.C., (1999) Empirical study of particle swarm optimization. In: *Proceedings of the 1999 Congress on Evolutionary Computation*. p. 1945–1950.
- Suneel, M., (2006) Chaotic sequences for secure CDMA. *Ramanujan Institute for Advanced Study in Mathematics*. p. 1–4.
- Wong, K., Man, K.P., Li, S., Liao, X., (2005) A more secure chaotic cryptographic scheme based on dynamic look-up table. *Circuits, Systems, and Signal Processing*, 24(5), 571–584.
- Wu, Q., Yan, H.S., Yang, H.B., (2008) A forecasting model based support vector machine and particle swarm optimization. In: *Proceedings of the 2008 workshop on power electronics and intelligent transportation system*. p. 218-222.
- Yuan, S.F., Chu, F.L., (2007) Fault diagnostics based on particle swarm optimization and support vector machines. *Mechanical Systems and Signal Processing*, 21(4), 1787–1798.
- Zhao, L., Yang, Y., (2009) PSO-based single multiplicative neuron model for time series prediction. *Expert Systems with Applications*, 36(2), 2805–2812.