

# Development of the Meta-Heuristic of PSOGA with K-means Algorithm

Y. Farhang

Faculty of Computer , Khoy Branch, Islamic Azad University, Khoy, Iran

## Abstract:

In this paper, a meta-heuristic approach was proposed for the hybridization of the K-means algorithm scheme. It obtained better results by developing a hybrid Genetic Algorithm-K-means (GA-K-means) and a hybrid Partial Swarm Optimization-K-means (PSO-K-means) method. In recent years, combinational optimization issues are introduced as critical problems in clustering algorithms to partition data in a way that optimizes the performance of clustering. K-means algorithm is one of the famous and more popular clustering algorithms which can be simply implemented and it can easily solve the optimization issue with less extra information. But the problems associated with K-means algorithm are high error rate, high intra cluster distance and low accuracy. In this regard, researchers have worked to improve the problem computationally, creating efficient solutions that lead to better data analysis through the K-means clustering algorithm. The aim of this study is to improve the accuracy of the K-means algorithm using hybrid and meta-heuristic methods. Finally, the meta-heuristic of Genetic Algorithm-Partial Swarm Optimization (GAPSO) and Partial Swarm Optimization-Genetic Algorithm (PSOGA) through the K-means algorithm were proposed. The approach adopted in this study successfully increased the accuracy rate of the clustering analysis and decreased its error rate and intra-cluster distance.

## Keywords:

*Hybrid Genetic Algorithm, K-means algorithm, Genetic Algorithm-Partial Swarm Optimization*

## 1. Introduction

K-means clustering, originating from signal processing is a method of vector quantization (Al-Jarrah et al., 2015). This is commonly applied to cluster analysis in data mining. The aim of K-means clustering is partitioning  $n$  observations into  $K$  clusters; in this case, each observation belongs to the cluster that has the nearest mean, which serves as a cluster's prototype (Xu and Wunsch, 2005, Dix, 2009, Jain, 2010). The problem has been proved to an NP-hard problem, though a number of efficient heuristic algorithms that have been proposed, which quickly converge to a local optimum. Generally, such algorithms are similar to the expectation-maximization algorithm for mixtures of Gaussian distributions through an iterative refinement approach that is adopted by both algorithms. In addition, both algorithms employ cluster centers for modeling the data. Nevertheless, in the expectation-

maximization mechanism, clusters are allowed to have various shapes, whereas K-means clustering usually finds clusters of similar spatial extent (Xu and Wunsch, 2005, Celebi et al., 2013). In the K-means clustering algorithms, there are a number of shortages and defects that should be improved (Afroozeh et al., 2012a, A. Afroozeh 2014, A. Afroozeh, 2014).

There are different methods to enhance and improve K-means clustering algorithm. One of these methods is to use the optimization method, in which a best element is selected from some of the set of available alternatives. Two important areas pertaining to optimization methods are the hybrid approach and the meta-heuristic approach (Akbari et al., 2016, Amiri et al., 2015, Afroozeh et al., 2014, Afroozeh et al., 2010, Afroozeh et al., 2015, Afroozeh et al., 2012b).

## Meta-Heuristic Method for Clustering

Heuristic is a technique applied to solving a problem more quickly compared to the use of classic methods, or finding an approximate solution in cases where classic methods have failed to propose any exact solution. This can be obtained through trading optimality, accuracy, completeness, or precision for speed. The heuristic can be considered as a shortcut for solving problems (Renner and Ekárt, 2003, Mohtashami et al., 2015).

On the other hand, a meta-heuristic is a heuristic of a higher level that is used for finding, generating, or selecting a lower-level heuristic or procedure (partial search algorithm), which might suggest an efficient solution to an optimization problem, in particular with limited computation capacity or imperfect information in computer science and mathematical optimization (Blum and Roli, 2003, Bianchi et al., 2009, Mladenović et al., 2007, Blum et al., 2011). In meta-heuristics, there may be few assumptions regarding the optimization problem being solved; thus they can be applied to various problems. In comparison with the iterative methods and optimization algorithms, meta-heuristics cannot guarantee a globally optimal solution for some classes of problems (Blum and Roli, 2003). In several meta-heuristics, some forms of stochastic optimization are implemented in such a way that the found solution depends on the set of generated random variables (Bianchi et al., 2009). Through searching over a

large set of feasible solutions, often meta-heuristics are able to find suitable solutions with less computational efforts compared to iterative methods, algorithms, or simple heuristics (Blum et al., 2011). This way, they can be considered as promising approaches to the optimization problems (Blum et al., 2011, Bianchi et al., 2009). In general, if two different algorithms are combined for solving problem, the approach is called hybrid approach. But if more than two algorithms or several heuristic algorithms are combined for this purpose, the approach is called a meta-heuristic. Note that hybrid of GA algorithm and K-means clustering algorithm has advantages for good clustering, and a hybrid of PSO Algorithm and K-means clustering algorithm has other advantages. It can be combined with the above-mentioned new methods in order to gain an algorithm that combines the advantages of both algorithms. It is called a meta-heuristic approach that is more successful than the previous method in clustering data.

### The Proposed I-PSO-K-means Algorithm

Due to the characteristics of K-means clustering algorithms, they can be combined and hybridized with many other algorithms. An optimization algorithm that can be combined with clustering algorithms is particle swarm optimization algorithm. Since, for solving the problem, the particle swarm optimization algorithm does not require additional information and labeling, the same as the K-means clustering algorithms, it can use this advantage to combine two algorithms. Furthermore, hybridization of two algorithms can help to solve one of clustering problems. This problem is that K-means clustering has high intra-cluster distance. Using Particle Swarm Optimization K-means (PSO-K-means), this intra-cluster distance can be reduced. The PSO-K-means algorithm is a hybrid algorithm explained in Chapter 2. For reducing

intra-cluster distance in the PSO-K-means, the Improved Particle Swarm Optimization-K-means (I-PSO-K-means) is proposed. This algorithm is fully described in the two following sections.

### Modeling of I-PSO-K-means Algorithm

This section improves the Particle Swarm Optimization Algorithm in K-means algorithm. Additionally, this section addresses the second objective of the study. One of the shortcomings of the PSO-K-means clustering algorithm is the high intra-cluster distance in the clustering of datasets, which can be low. To this end, the Improved Particle Swarm Optimization-K-means algorithm (I-PSO-K-means) is proposed. In the following, the design of I-PSO-K-means algorithm is described. The proposed algorithm in this section comprises eight important steps: initialization, compare for obtaining Pbest, compare for obtaining Gbest, calculating the function, checking the Max-domain, checking the Min-domain, and checking the repeat and running K-means. These steps are shown in Figure 4.5.

### Implementation of the I-PSO-K-means Algorithm

In this section, the implementation of the I-PSO-K-means clustering algorithm (Improved Particle Swarm Optimization-K-means algorithm) is elaborated. As mentioned in the previous section, the algorithm proposed here is a hybrid of the PSO algorithm and the K-means clustering algorithm. In the following, the implementation of the I-PSO-K-means clustering algorithm is described. The proposed algorithm in this section has fourteen main steps. These steps are shown in Figure 1.

Improved Particle Swarm Optimization-K-means	
1.	Start.
2.	Selection the Initial Population.
	a. Applying Dataset
	b. Finding Domain of All Attributes in Dataset ( $h_1, h_2, \dots$ )
	i. Finding Minimum Domain ( $h_1\text{-min}, h_2\text{-min}, \dots$ )
	ii. Finding Minimum Domain ( $h_1\text{-max}, h_2\text{-max}, \dots$ )
	iii. Finding Domain ( $h_1 = h_1\text{-max} - h_1\text{-min}, h_2 = h_1\text{-max} - h_2\text{-min}, \dots$ )
	iv. For $f=1$ to 50
	1. Finding 50 Cluster Centers by Randomly in Domain (Initial Population)
	2. Running K-means Algorithm on Initial Population
	a. For $n=1$ to 50
	i. Mining of Features $n^{\text{th}}$ Cluster Center from Dataset ( $m_1, m_2, \dots$ )
	ii. For $K=1$ to $N$ /* $N$ is No. members in Dataset. /*
	1. Mining of Features $K^{\text{th}}$ Row from Dataset ( $a_1, a_2, \dots$ )
	2. Calculate Distance ( $m_1, m_2, \dots$ ) and ( $a_1, a_2, \dots$ ) by Euclidean
	3. Finding of Minimum Distance between Cluster Centers
	4. Placement in Cluster that it has Minimum Distance
	5. Calculate total of distance ( $S_1, S_2, \dots$ )
	iii. end For
	iv. Calculate $S$ ( $S = S_1 + S_2 + \dots$ ) /* $S$ is Intra-cluster Distance. /*
	v. Placement Cluster Centers and $S$ in the Matrix. (Matrix Name is $P$ )
	b. end For
3.	Evaluation
	a. Sorting Rows of Matrix $P$ Based on $S$ (Descending $P$ is $P_1$ )

```

4. Selection
    a. Selection one members from  $P_1$  for Next Section. (Lowest  $S$ )
5. Placement of Previously member in  $P_0$  ( $P_0(f,1:Co)$ )
v. end For
3. Evaluation
    a. Placement  $P_0$  in  $P$  ( $P$  is 50 members with  $S$  minimum)
    b. Sorting Rows of Matrix  $P$  Based on  $S$  (Descending  $P$  is  $P_1$ )
4. Selection
    a. Selection first members from  $P_1$  for  $Gbest$ . (Lowest  $S$ )
5. Initial Values
    a.  $Gbest = P_1(1,1:Co)$ ;
    b.  $W = 0.7299$ ;
    c.  $C_1 = 1.4963$ ;
    d.  $C_2 = 1.4963$ ;
    e.  $R_1 = \text{Random}(0 \text{ to } 1)$ 
    f.  $R_2 = \text{Random}(0 \text{ to } 1)$ 
6. Specified  $Pbest$ 
    a.  $Pbest = P(1,1:Co)$ ;
7. Specified  $Xt$ 
    a.  $Xt = P(1,1:Co)$ ;
8.  $Vt = 0$ ;
9. PSO Operator
    a. For  $f = 1$  to 50
        i.  $Vtt = (W * Vt) + (C1 * R1)(Pbest - Xt) + (C2 * R2)(Gbest - Xt)$ 
        ii.  $Xtt = Xt + Vtt$ 
        iii. Checking  $Xtt$ 
            1. If ( $Xtt > \text{Max\_Max\_Domain}$ )
                a.  $Xtt = \text{Max\_Domain}$ 
            2. If ( $\text{Min\_Domain} > Xtt$ )
                a.  $Xtt = \text{Min\_Domain}$ 
            3. Checking for All Columns
        iv.  $Xt = Xtt$ ;
        v.  $Vt = Vtt$ 
        vi.  $S = 0$ ; ( $S = 0, S_1 = 0, S_2 = 0, \dots$ )
        vii. Placement  $Xtt$  on  $P_2$ 
        viii. Running K-means Algorithm on  $Xtt$ 
            1. For  $K = 1$  to  $N$  /*  $N$  is number of members in Dataset. */
                a. Mining of Features  $K^{\text{th}}$  Row from Dataset ( $a_1, a_2, \dots$ )
                b. Calculate Distance ( $m_1, m_2, \dots$ ) and ( $a_1, a_2, \dots$ ) by Euclidean
                c. Finding of Minimum Distance between Cluster Centers
                d. Placement in Cluster that it has Minimum Distance
                e. Calculate total of distance ( $S_1, S_2, \dots$ )
            2. end For
            3. Calculate  $S$  ( $S = S_1 + S_2 + \dots$ ) /*  $S$  is Intra-cluster Distance. */
            4. Placement Cluster Centers and  $S$  in the Matrix. (Matrix Name is  $P_3$ )
            5. Evaluation
                a. Sorting Rows of Matrix  $P_3$  Based on  $S$  (Descending  $P_3$  is  $P_4$ )
            6. Selection
                a. Selection 50 members from  $P_4$  for Next Section. (Lowest  $S$ )
        ix. Selection
            1. Placement Cluster Centers with intra-cluster centers  $P_1$  on  $P_5$  (50)
            2. Placement Cluster Centers with intra-cluster centers  $P_3$  on  $P_5$  (50)
        x. Evaluation
            1. Sorting Rows of Matrix  $P_5$  Based on  $S$  (Descending  $P_5$  is  $P_6$ )
        xi. Specified  $Pbest$ 
            1.  $Pbest = P_4(1,1:Co)$ ;
        xii. Specified  $Gbest$ 
            1.  $Gbest = P_6(1,1:Co)$ ;
    b. end For
10. Iteration of Steps 6 to 9 (50 times).
11. Selection
    a. Selection first members from  $P_6$  for Final Answer.
12. Running K-means Algorithm on Final Answer
    a. For  $K = 1$  to  $N$  /*  $N$  is number of members in Dataset. */
        i. Mining of Features  $K^{\text{th}}$  Row from Dataset ( $a_1, a_2, \dots$ )
        ii. Calculate Distance ( $m_1, m_2, \dots$ ) and ( $a_1, a_2, \dots$ ) by Euclidean
        iii. Finding of Minimum Distance between Cluster Centers
        iv. Placement in Cluster that it has Minimum Distance
        v. Calculate total of distance ( $S_1, S_2, \dots$ )
    b. end For
    c. Calculate  $S$  ( $S = S_1 + S_2 + \dots$ ) /*  $S$  is Intra-cluster Distance. */
13. Drawing of Chart
14. Stop.

```

Figure 1. The Pseudo Code of I-PSO-K-means Clustering Algorithm

In the I-PSO-K-means clustering algorithm, there is innovation in different parts of the algorithm. This algorithm is a hybrid of the K-means clustering algorithm and the particle swarm optimization algorithm, which reduces the intra-cluster distance in the K-means clustering algorithm. In the next section, the first proposed algorithm is investigated using different datasets and the results are compared with those of other algorithms.

### Analysis of I-PSO-K-means Algorithm

Here, the analysis of the results obtained from the I-PSO-K-means clustering algorithm is presented. The I-PSO-K-means algorithm is related to the second phase of this study, namely intra-cluster distance. For intra-cluster distance, four criteria are taken into consideration the best of intra-cluster distance, worst of intra-cluster distance, average of intra-cluster distance, and standard deviation of intra-cluster distance. To better assess the performance of the proposed algorithm (I-PSO-K-means clustering

algorithm); the K-means algorithm and PSO-K-means clustering algorithm are examined. For the K-means algorithm, the algorithm proposed by (Meilă, 2006) is selected because this article maintains the framework of the K-means algorithm. For the Particle Swarm Optimization-K-means algorithm, the algorithm of (Tsai and Kao, 2011) is chosen because the intra-cluster distance factor for evaluation in this instance is similar to the second phase. In this section, the proposed algorithm (the I-PSO-K-means algorithm), the PSO-K-means algorithm, and three previous algorithms have been tested using 6 data sets (Balance, Blood, Breast, Iris, Pima, and Wine).

In Table 1, the data and results of six data are expressed for five algorithms. To better assess the performance of the algorithms, the algorithms are run 20 times. The results are average of intra-cluster distance, standard deviation of intra-cluster distance, best of intra-cluster distance and worst of intra-cluster distance for all algorithms.

Table 1: The results I-PSO-K-means algorithm for 20 times running.

Dataset	Name of Algorithm	Intra-cluster distance			
		Best	Worst	Average	Std. Dev.
Balance	K-means	1426.21	1439.08	1432.57	4.29
	GA-K-means	1423.55	1429.85	1427.31	2.10
	I-GA-K-means	1423.35	1429.85	1426.44	1.94
	PSO-K-means	1423.91	1429.85	1426.59	1.68
	I-PSO-K-means	<b>1423.25</b>	<b>1426.8</b>	<b>1425.14</b>	<b>1.00</b>
Blood	K-means	469637	490933	474810	8310
	GA-K-means	409325	429637	414056	5447
	I-GA-K-means	409011	419289	413217	2679
	PSO-K-means	405538	419875	413251	3951
	I-PSO-K-means	<b>408011</b>	<b>411561</b>	<b>409501</b>	<b>959</b>
Breast	K-means	3056.96	3095.95	3067.03	15.50
	GA-K-means	3054.65	3091.05	3066.83	10.63
	I-GA-K-means	3051.09	3081.61	3061.39	7.85
	PSO-K-means	3052.16	3079.76	3060.98	9.16
	I-PSO-K-means	<b>3050.30</b>	<b>3077.3</b>	<b>3058.8</b>	<b>8.67</b>
Iris	K-means	97.32	122.47	102.86	9.55
	GA-K-means	97.03	99.18	97.86	0.67
	I-GA-K-means	96.69	98.73	97.41	0.56
	PSO-K-means	96.11	98.05	96.67	0.54
	I-PSO-K-means	<b>96.04</b>	<b>98.01</b>	<b>96.63</b>	<b>0.55</b>
Pima	K-means	52867	52072	52194	238
	GA-K-means	48512	59348	52141	2306
	I-GA-K-means	47936	55170	51550	2374
	PSO-K-means	48033	58918	58918	2550
	I-PSO-K-means	<b>47267</b>	<b>55654</b>	<b>51489</b>	<b>2403</b>
Wine	K-means	16555	18467	16811	496
	GA-K-means	16324	17294	16536	229
	I-GA-K-means	16292	16466	16382	53
	PSO-K-means	16288	16316	16299	8.18
	I-PSO-K-means	<b>16284</b>	<b>16315</b>	<b>16296</b>	<b>6.63</b>

In the above table can be seen that the proposed algorithm has better performance. In the next section, it is discussed results of this table.

### Discussion of I-PSO-K-means Algorithm

In this part, the results of I-PSO-K-means clustering algorithm are discussed. Since the I-PSO-K-means algorithm is related to the second phase of the study; in this

phase, the comparison factor is intra-cluster distance. Therefore, in this section, two important areas, namely the intra-cluster distance that is the average of intra-cluster distance and the standard deviation of intra-cluster

distance are analyzed. In Figure 2, the average of intra-cluster distance is shown for 20 times of running of the five algorithms.

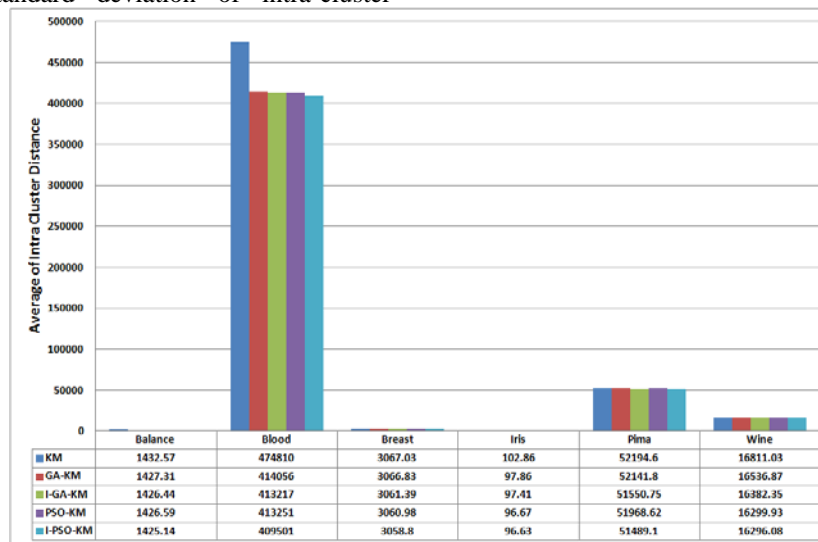


Figure 2: The average of intra-cluster distance in I-PSO-KM

In the above figure, it can be seen that the average of intra-cluster distance in the proposed algorithm (Improved Particle Swarm Optimization-K-means) is better than the previous algorithms. Therefore, the performance of the proposed algorithm in this phase can be better than previous algorithms.

In Figure 3, the standard deviation of intra-cluster distance is shown for 20 times that the five algorithms ran. Lower standard deviation in algorithm shows that the algorithm is more stable.

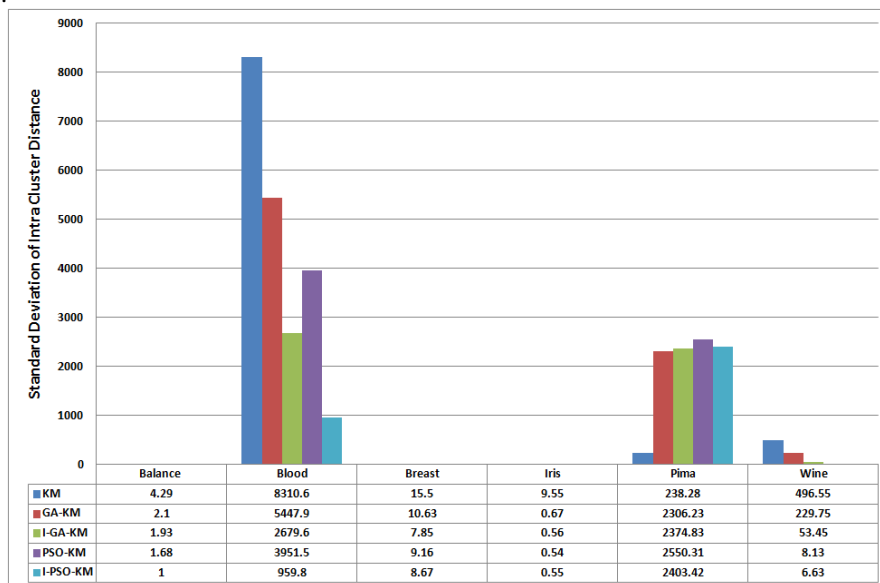


Figure 3: The standard deviation of intra-cluster distance in I-PSO-KM

In the Figure 4, it can be observed that the standard deviation of intra-cluster distance in the proposed algorithm (Improved Particle Swarm Optimization -K-

means) is better than the previous algorithm. Therefore, the proposed algorithm in this phase has a better performance compared to previous algorithms.

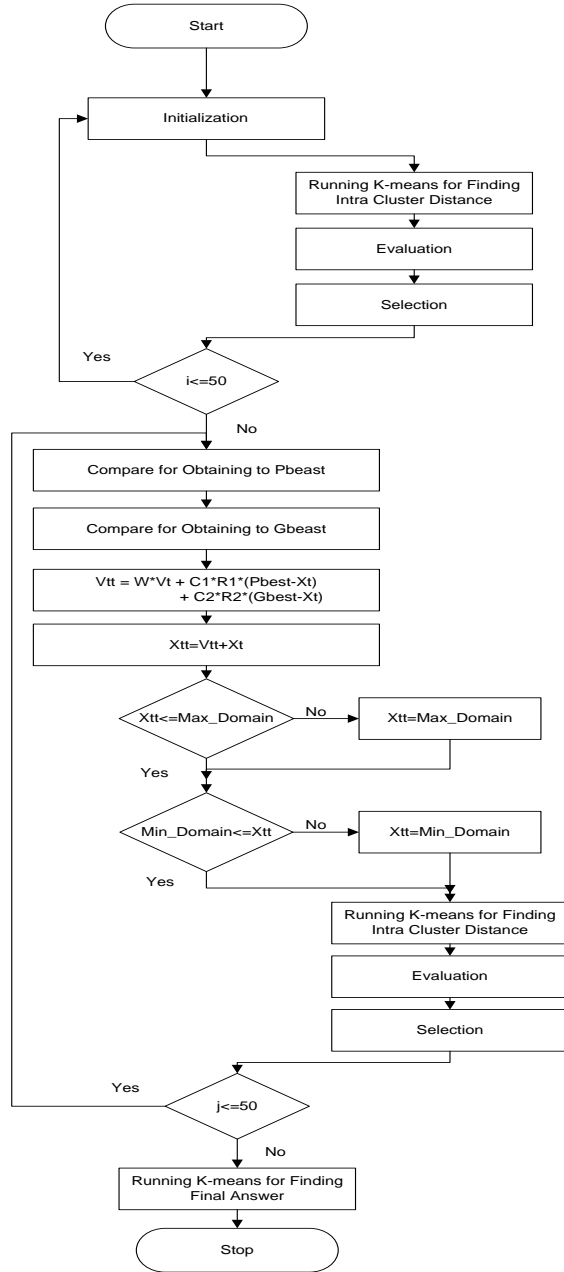


Figure 4: The Flowchart of I-PSO-K-means Algorithm

First, the number of initial population is chosen by random for cluster centers. This step is done for 40 times in order to obtain acceptable initial values. This phase is one of the new features in the I-PSO-K-means algorithm. After various experiments on various datasets, it was found that the selection of initial value of the PSO algorithm was very important. If the initial value is selected correctly, final result can be reached quickly and it reduces the intra-cluster distance by selecting appropriate initial value. Second, obtaining Pbest for function is addressed. The Pbest is the best value among local values, which is

selected to compare local values. All local values are calculated for intra-cluster distance by K-means clustering algorithm. After calculating intra-cluster distance, the minimum intra-cluster distance is selected (best value) for new Pbest to be used in the next step.

The third step addresses the achievement of Gbest for function. The Gbest is the best value among global values, which is selected to compare local values. All global values are calculated for intra-cluster distance by K-means clustering algorithm. After the calculation of the intra-cluster distance, the minimum intra-cluster distance is chosen (the best value) for new Gbest to be used in the next step.

Fourth, the function Vtt is calculated, which has three important items. The first item moves toward previously line, the second one move toward local best line, and the third one moves toward global best line. Vtt should movetoward goal function. Then, the Xtt is calculated as Vtt and Xt (previously Xtt).

The fifth and sixth steps are checking Xtt in the Max\_domain and checking Xtt in the Min\_domain. It should be Xtt into domain because if it is not, the result cannot be close to the goal function. In the I-PSO-K-means algorithm, new method is used for checking domain of Xtt. If Xtt > Max\_domain, then Xtt = Max\_domain, and if Xtt < Min\_domain, then Xtt = Min\_domain. Additionally, by selecting an appropriate domain, the intra-cluster distance can be reduced.

Figure 5, demonstrates an example of new Checking domain Xtt in the I-PSO-K-means clustering algorithm. In I-PSO-K-means algorithm, for checking domain, Xtt is created in two models. First, if Xtt > Max\_domain, then Xtt = Max\_domain. Second, if Xtt < Min\_domain, then Xtt = Min\_domain. In this condition, the answer is close to the goal function.

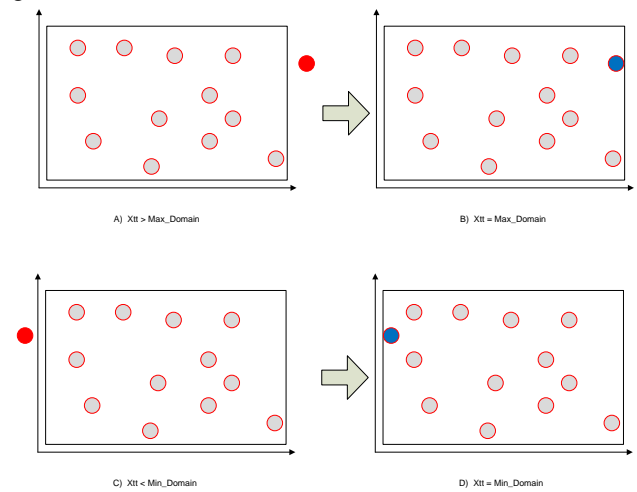


Figure 5: Checking of the domain Xtt in the I-PSO-K-means Algorithm

Seventh, the repetition of the main part of algorithm is checked. The main part of I-PSO-K-means algorithm is run 50 times. This is because the algorithm reaches the balance after 50 times running.

The eighth and final stage in the I-PSO-K-means algorithm is running the K-means clustering algorithm signal time. After running the initial step and the main step for finding final answer, the K-means is run to obtain data for analysis.

## Conclusion

For the general selection, the initial population (50 members) and the calculated function population (50 members) are combined, and then the collection of the population is evaluated and sorted, the extra population is removed, and the new initial population is selected for the next iteration to reach the goal function. All these steps are repeated 50 times to reach the optimal result (cluster centers). Finally, optimal solution to the K-means clustering algorithm is used to obtain the results of the analysis. In the next section, the implementation of I-PSO-K-means clustering algorithm will be described in detail.

## References

- [1] AFROOZEH, A. Z., S. E. POURMAND AND I. S. AMIRI 2014. Attosecond pulse generation using nano ring waveguides. *ijcls*, 5, 5.
- [2] AFROOZEH, A. Z., S. E. POURMAND AND I. S. AMIRI 2014. Determination of Suitable Material to Control of Light. *IJBPAS*, 3, 11.
- [3] AFROOZEH, A., AMIRI, I., BAHADORAN, M., ALI, J. & YUPAPIN, P. 2012a. Simulation of soliton amplification in micro ring resonator for optical communication. *Jurnal Teknologi*, 55, 271-277.
- [4] AFROOZEH, A., AMIRI, I., CHAUDHARY, K., ALI, J. & YUPAPIN, P. 2014. Analysis of optical ring resonator. *Advances in Laser and Optics Research*, 1-16.
- [5] AFROOZEH, A., AMIRI, I., KOUHNAVARD, M., BAHADORAN, M., JALIL, M., ALI, J. & YUPAPIN, P. 2010. Optical memory time using multi bright soliton. *International conference on experimental mechanics (ICEM)*, Kuala Lumpur, Malaysia, 29.
- [6] AFROOZEH, A., AMIRI, I., POURMAND, S., ZEINALINEZHAD, A., ALAVI, S. & AHMAD, H. 2015. Comparison of Control Light Using Kramers-Kronig Method by Three Waveguides. *Journal of Computational and Theoretical Nanoscience*, 12, 1864-1868.
- [7] AFROOZEH, A., AMIRI, I., SAMAVATI, A., ALI, J. & YUPAPIN, P. THz frequency generation using MRRs for THz imaging. *Enabling Science and Nanotechnology (ESciNano)*, 2012 International Conference on, 2012b. *IEEE*, 1-2.
- [8] AKBARI, E., BUNTAT, Z., NIKOUKAR, A., KHEIRANDISH, A., KHALEDIAN, M. & AFROOZEH, A. 2016. Sensor application in Direct Methanol Fuel Cells (DMFCs). *Renewable and Sustainable Energy Reviews*, 60, 1125-1139.
- [9] AL-JARRAH, O. Y., YOO, P. D., MUHAIDAT, S., KARAGIANNIDIS, G. K. & TAHA, K. 2015. Efficient Machine Learning for Big Data: A Review. *Big Data Research*.
- [10] AMIRI, I., ALAVI, S., BAHADORAN, M., AFROOZEH, A. & AHMAD, H. 2015. Nanometer bandwidth soliton generation and experimental transmission within nonlinear fiber optics using an add-drop filter system. *Journal of Computational and Theoretical Nanoscience*, 12, 221-225.
- [11] BIANCHI, L., DORIGO, M., GAMBARDELLA, L. M. & GUTJAHR, W. J. 2009. A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing: an international journal*, 8, 239-287.
- [12] BLUM, C., PUCHINGER, J., RAIDL, G. R. & ROLI, A. 2011. Hybrid metaheuristics in combinatorial optimization: A survey. *Applied Soft Computing*, 11, 4135-4151.
- [13] BLUM, C. & ROLI, A. 2003. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM computing surveys (CSUR)*, 35, 268-308.
- [14] CELEBI, M. E., KINGRAVI, H. A. & VELA, P. A. 2013. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, 40, 200-210.
- [15] DIX, A. 2009. *Human-computer interaction*, Springer.
- [16] JAIN, A. K. 2010. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31, 651-666.
- [17] MEILĀ, M. The uniqueness of a good optimum for k-means. *Proceedings of the 23rd international conference on Machine learning*, 2006. *ACM*, 625-632.
- [18] MLADENOVIĆ, N., BRIMBERG, J., HANSEN, P. & MORENO-PÉREZ, J. A. 2007. The  $p$ -median problem: A survey of metaheuristic approaches. *European Journal of Operational Research*, 179, 927-939.
- [19] MOHTASHAMI, A., TAVANA, M., SANTOS-ARTEAGA, F. J. & FALLAHIAN-NAJAFABADI, A. 2015. A novel multi-objective meta-heuristic model for solving cross-docking scheduling problems. *Applied Soft Computing*, 31, 30-47.
- [20] RENNER, G. & EKÁRT, A. 2003. Genetic algorithms in computer aided design. *Computer-Aided Design*, 35, 709-726.
- [21] TSAI, C.-Y. & KAO, I.-W. 2011. Particle swarm optimization with selective particle regeneration for data clustering. *Expert Systems with Applications*, 38, 6565-6576.
- [22] XU, R. & WUNSCH, D. 2005. Survey of clustering algorithms. *Neural Networks, IEEE Transactions on*, 16, 645-678.