Swarm Intelligence

By

Nasser M.

Outline

- Swarm Intelligence
- Metahuristics
- Particle Swarm Optimization (PSO)
- Ant Colony Optimization (ACO)
- Case Study: Data Clustering Using PSO
- Conclusion

Swarm Intelligence

"The emergent collective intelligence of groups of simple agents."

(Bonabeau et al, 1999)

Characteristics of Swarms

- Composed of many individuals
- Individuals are homogeneous
- Local interaction based on simple rules
- Self-organization (No centralized Control)

Swarm Intelligence Algorithms

- **Particle Swarm Optimization (PSO)**
- **Ant Colony Optimization**
- Artificial Bee Colony Algorithm
- Artificial Immune Systems Algorithm

Search Techniques

- Deterministic Search Techniques
  - Branch and Bound
  - Steepest Descent
  - Newton-Raphson
  - Simplex based Technique
- Stochastic or Random Search Techniques
  - Swarm Intelligence
  - Genetic Algorithm
  - Differential Evolution
  - Simulated Annealing

Components of Search Techniques

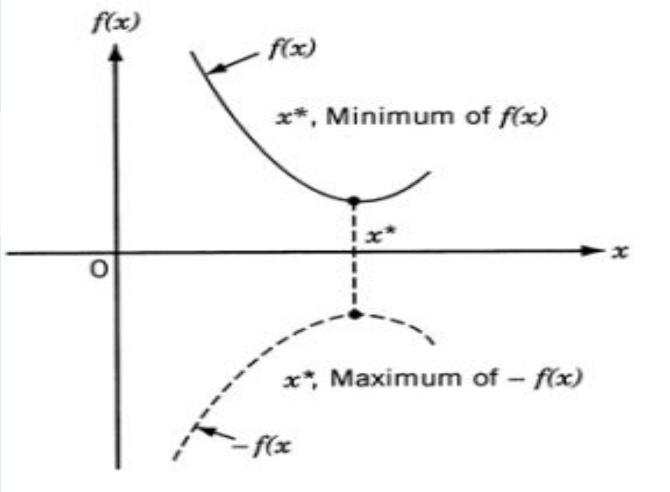
- Initial solution
- Search direction
- Update criteria
- Stopping criteria
- All above elements can be either
  - Deterministic or Stochastic
  - Single points or population based

Heuristics

- Heuristic (or approximate) algorithms aim to find a good solution to a problem in a reasonable amount of computation time – but with no guarantee of "goodness" or "efficiency" (cf. exact or complete algorithms).

- Heuristic is used to solve NP-Complete Problem , a class of decision problem.

  Every decision problem has equivalent optimization problem



Minimum of $f(x)$ is same as maximum of $-f(x)$.

Metaheuristics

- Metaheuristics are (roughly) high-level strategies that combining lower-level techniques for exploration and exploitation of the search space.

- Metaheristcs refers to algorithms including

  -Evolutionary Algorithms

  Swarm Intelligence , Genetic Algorithm , Differential Evolution , Evolutionary programing

  - Simulated Annealing

  - Tabu Search,

Fundamental Properties of Metaheuristics

- Metaheuristics are strategies that “guide” the search process.
- The goal is to efficiently explore the search space in order to find (near-)optimal solutions.
- Metaheuristic algorithms are approximate and usually non-deterministic.
- Metaheuristics are not problem-specific.

Outline

- Swarm Intelligence
- Metahuristics
- Particle Swarm Optimization (PSO)
- Ant Colony Optimization (ACO)
- Case Study: Data Clustering Using PSO
- Conclusion

Particle Swarm Optimization

Source https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSKhLOQonxCPbIKl6GE0UWhaJJByuKpYFtWDUovH1Ss0HUaaWcq

Particle Swarm Optimization (PSO)

- PSO is stochastic optimization technique proposed by Kennedy and Eberhart (1995) [2].

- A population based search method with position of particle is representing solution and Swarm of particles as searching agent.

- PSO is a robust evolutionary optimization technique based on the movement and intelligence of swarms.

- PSO find the minimum value for the function.

Particle Swarm Optimization (PSO)

- The idea is similar to bird flocks searching for food.
  - Bird = a particle, Food = a solution
  - *pbest* = the best solution (fitness) a particle has achieved so far.
  - *gbest* = the global best solution of all particles within the swarm

PSO Search Scheme

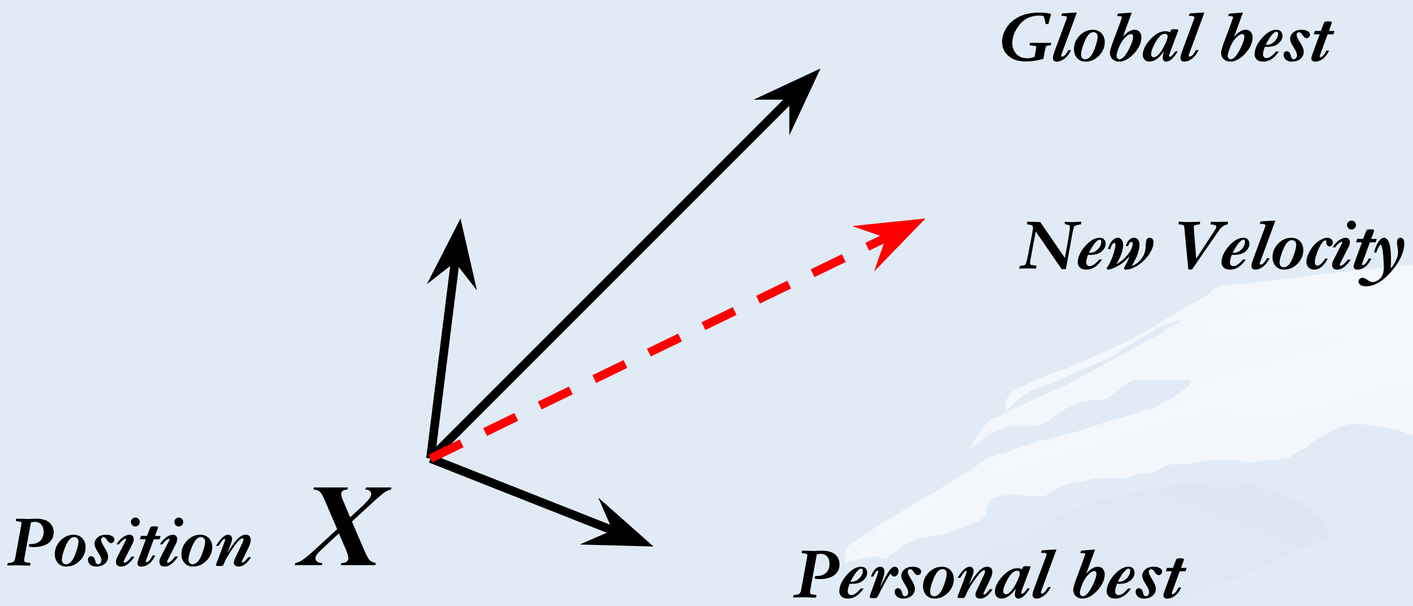
- **pbest** : the best solution achieved so far by that particle.
- **gbest** : the best value obtained so far by any particle in the neighborhood of that particle.

- The basic concept of PSO lies in accelerating each particle toward its pbest and the gbest locations, with a random weighted acceleration at each time.

PSO Search Scheme

- PSO uses a number of agents, i.e., particles, that constitute a swarm flying in the search space looking for the best solution.

- Each particle is treated as a point (candidate solution) in a N-dimensional space which adjusts its "flying" according to its own flying experience as well as the flying experience of other particles.


Global best
New Velocity
Position X
Personal best

Particle Swarm Optimization (PSO)

Each particle tries to modify its position X using the following formula:

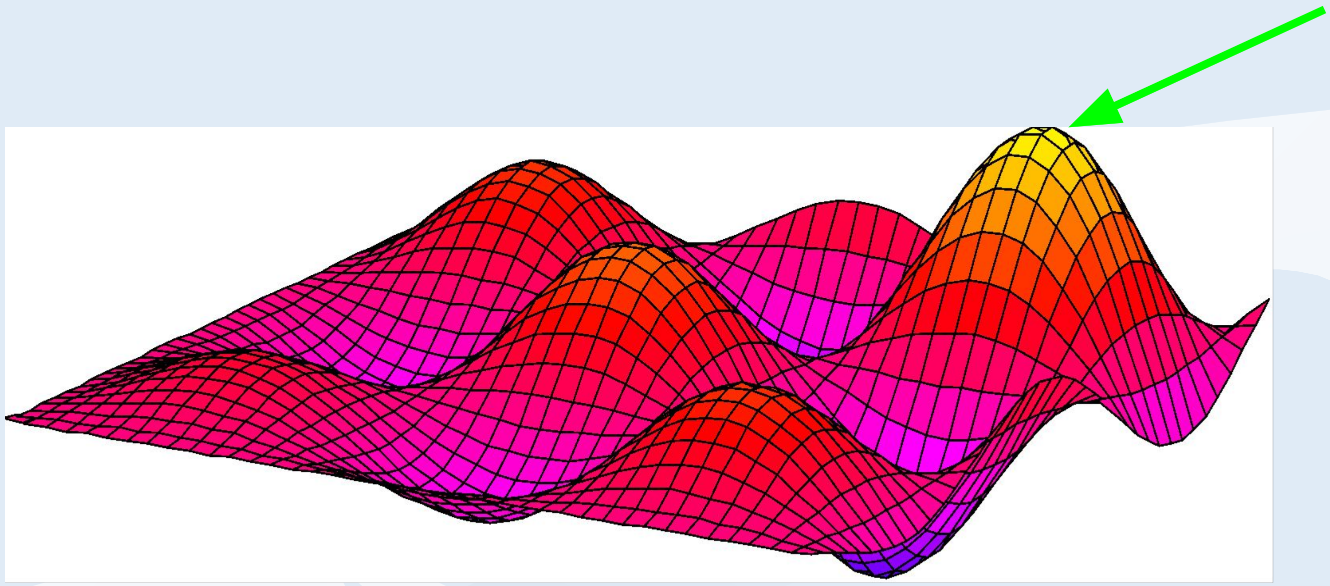
$$X(t+1) = X(t) + V(t+1) \quad (1)$$

$$V(t+1) = wV(t) + c_1 \times rand() \times (X_{pbest} - X(t)) + c_2 \times rand() \times (X_{gbest} - X(t)) \quad (2)$$

$V(t)$	velocity of the particle at time t
$X(t)$	Particle position at time t
w	Inertia weight
c_1, c_2	learning factor or accelerating factor
rand	uniformly distributed random number between 0 and 1
X_{pbest}	particle's best position
X_{gbest}	global best position

Alpine function

$$f(x_1,\ldots,x_D) = \sin(x_1)\ldots\sin(x_D)\sqrt{x_1\ldots x_D}$$



Particle fly and search for the highest peak in the search space

PSO Algorithm

The PSO algorithm pseudocode [2] as following:

Input: Randomly initialized position and velocity of Particles:
Xi (0) and Vi (0)

Output: Position of the approximate global minimum X^*

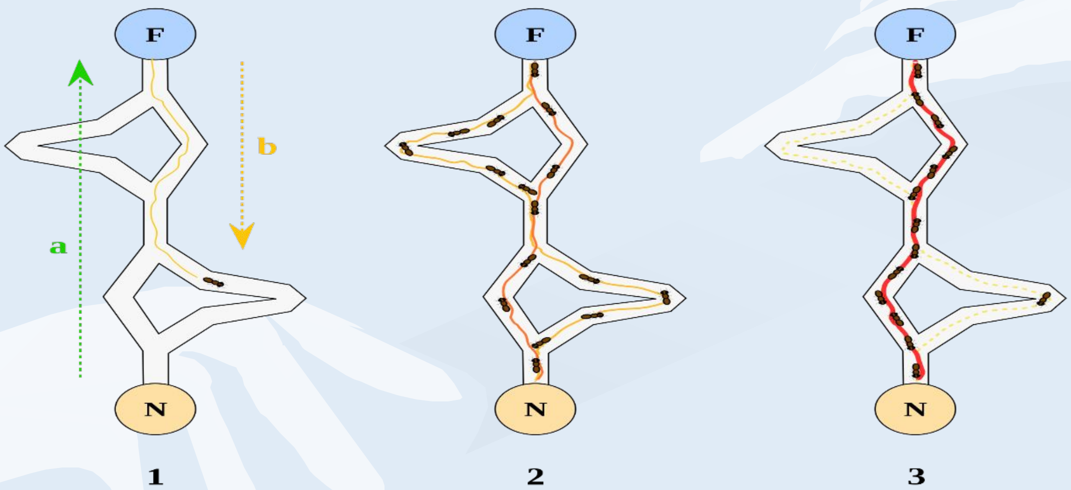
```
while terminating condition is not reached do
    for i = 1 to number of particles do
        Calculate the fitness function f
        Update personal best and global best of each particle
        Update velocity of the particle using Equation 2
        Update the position of the particle using equation 1
    end for
end while
```

Outline

- Swarm Intelligence
- Metahuristics
- Particle Swarm Optimization (PSO)
- Ant Colony Optimization (ACO)
- Case Study: Data Clustering Using PSO
- Conclusion

Ant Colony Optimization

"Ant Colony Optimization (ACO) studies artificial systems that take inspiration from the *behavior of real ant colonies* and which are used to solve discrete optimization problems." ACO Website [1]



Source: http://upload.wikimedia.org/wikipedia/commons/thumb/a/af/Aco_branches.svg/2000px-Aco_branches.svg.png

Ant Colony Optimization

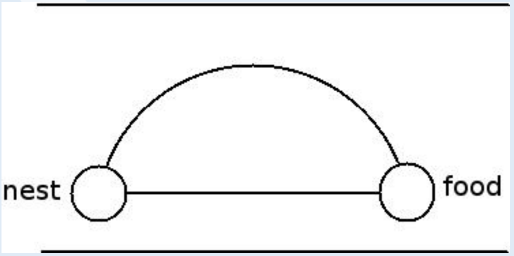
- Probalistic Techniques to solve optimization Problem
- It is a population based metaheuristic used to find approximate solution to an optimization problem.
- The Optimization Problem must be written in the form of path finding with a weighted graph

Application of ACO

- Shortest paths and routing
- Assignment problem
- Set Problem

Idea

- The way ants find their food in shortest path is interesting.
- Ants hide pheromones to remember their path.
- These pheromones evaporate with time.
- Whenever an ant finds food , it marks its return journey with pheromones.
- Pheromones evaporate faster on longer paths.



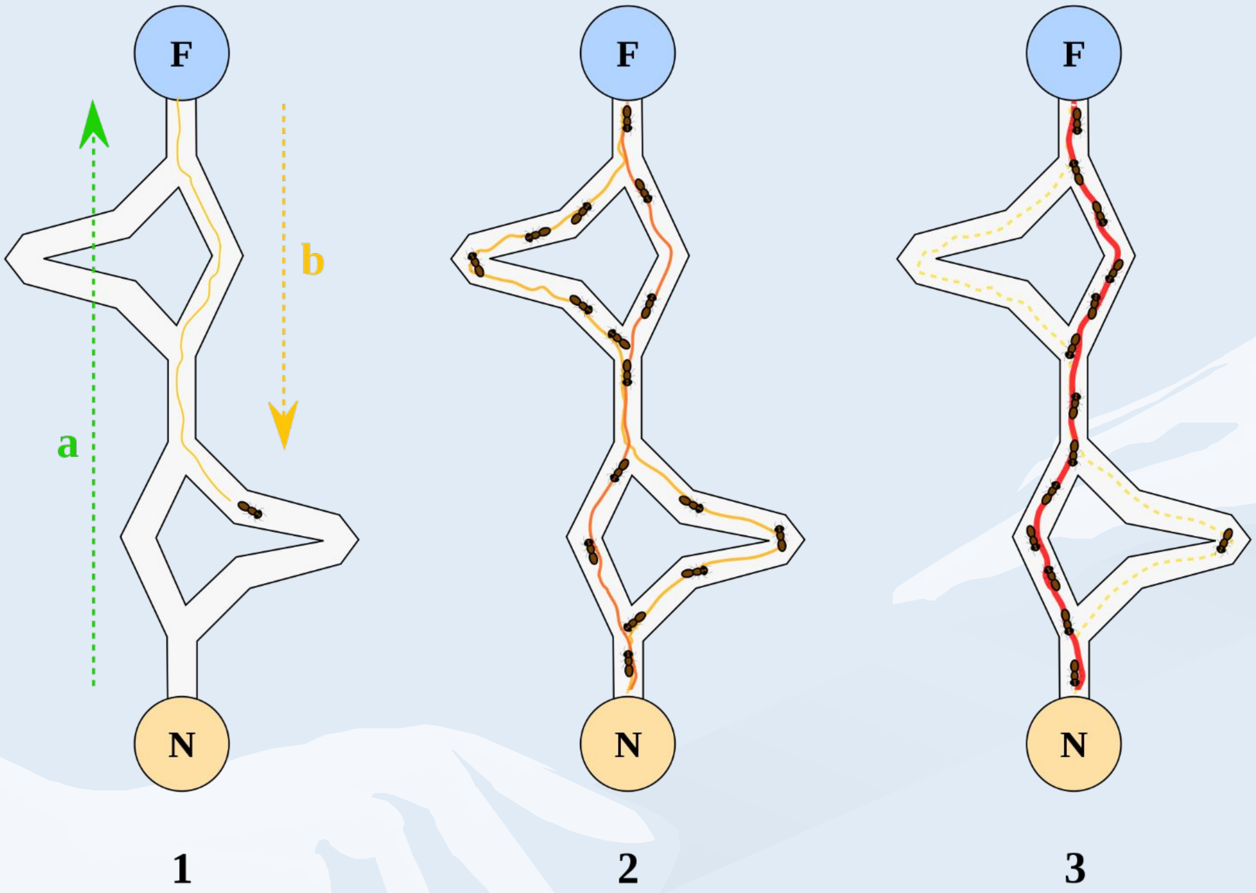
Idea (cont.)

- Shorter paths serve as the way to food for most of the other ants.
- The shorter path will be reinforced by the pheromones further.
- Finally , the ants arrive at the shortest path.

ACO Concept

- Ants navigate from nest to food source. Ants are blind!
- Shortest path is discovered via pheromone trails. Each ant moves at random
- Pheromone is deposited on path
- More pheromone on path increases probability of path being followed

Ant Colony Optimization



Source: http://upload.wikimedia.org/wikipedia/commons/thumb/a/af/Aco_branches.svg/2000px-Aco_branches.svg.png

Ant Colony Algorithm

Algorithm 1 The Ant Colony Optimization Metaheuristic

```
Set parameters, initialize pheromone trails
while termination condition not met do
    ConstructAntSolutions
    ApplyLocalSearch (optional)
    UpdatePheromones
end while
```

- ConstructAntSolutions: Partial solution extended by adding an edge based on stochastic and pheromone considerations.
- ApplyLocalSearch: problem-specific, used in state-of-art ACO algorithms.
- UpdatePheromones: increase pheromone of good solutions, decrease that of bad solutions (pheromone evaporation).

Outline

- Swarm Intelligence
- Metahuristics
- Particle Swarm Optimization (PSO)
- Ant Colony Optimization (ACO)
- Case Study: Data Clustering Using PSO
- Conclusion

Clustering

Clustering : partitioning of a data set into subsets - clusters, so that the data in each subset share some common features often based on some similarity.

Clustering is NP-hard problem

- No optimal solution in polynomial time
- We need heuristic efficient algorithm (approximation solution)
- can be formulated as optimization problem

Partitioning Clustering

- The partitioning techniques usually produce clusters by optimizing a criterion function

- In partitioning clustering ,the squared error criterion is minimized, which tends to work well with isolated and compact clusters [3].

$$\min \; f = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2$$

Where $\boldsymbol{x}_i$ data pattern belong to cluster *i and* $\boldsymbol{\mu}_j$ is the center of cluster *j* and *k* is number of clusters

Continue

Partitioning clustering algorithms such as Kmeans , Kmodes are relatively efficient but have several drawbacks.

drawbacks

- Often terminate at local minimum
- Generate empty clusters
- Unable to handle noisy data and *outliers*

Solution : Clustering using PSO algorithm

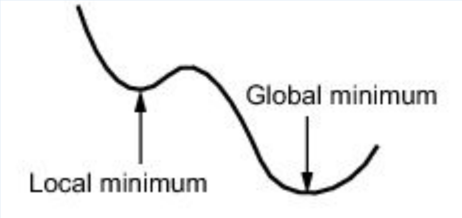
Idea: Using PSO algorithm to minimize the objective function of clustering (squared error criterion)

$$\min \ f = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2$$

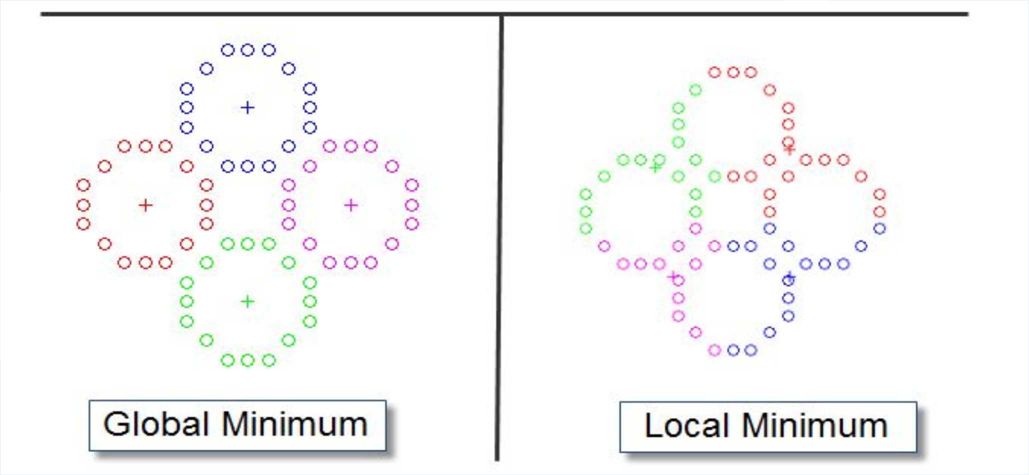
Continue

Why PSO based clustering

-Terminate at global optimum



- High quality than tradition methods such as Kmeans
- Not sensitive for noisy and outlier data
- Avoid problem of generating empty clusters



PSO Algorithm

The PSO algorithm peudocode [2] as following:

Input: Randomly initialized position and velocity of Particles:
$Xi\ (0)$ and $Vi\ (0)$

Output: Position of the approximate global minimum X^*

```
while terminating condition is not reached do
    for i = 1 to number of particles do
        Calculate the fitness function f
        Update personal best and global best of each particle
        Update velocity of the particle using Equation 2
        Update the position of the particle using equation 1
    end for
end while
```

Data Clustering Formulation

- The aim is to partition unlabeled data to k disjoint classes by optimizing a criterion function (square error function)

- This is achieved by optimizing the following objective

$$\min \ f = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2$$

Where $\boldsymbol{x}_i$ Where data pattern belong to cluster i

$\boldsymbol{\mu}_j$ is the center of cluster j and k is number of clusters

Result in

high <u>intra-class</u> similarity: maximize distances between clusters

low <u>inter-class</u> similarity: minimize distances within clusters

PSO Clustering Algorithm

- Each particle maintains a vector $\boldsymbol{Vi} = (\boldsymbol{C_1}, \boldsymbol{C_2}, \ldots, \boldsymbol{C_i}, \ldots, \boldsymbol{C_k})$, where $\boldsymbol{C_i}$ represents the ***i***th cluster centroid vector and k is the number of clusters.

- The particle adjusts the centroid vector' position in the vector space at each generation (iteration)
  $X(\mathrm{t}+1) = X(\mathrm{t}) + V(\mathrm{t}+1)$

-For example : suppose the k= 4 , and the particle *i* maintain vector $\boldsymbol{Vi} = \{(1,2), (3,5), (7,4), (8,2)\}$ at t =1
at t = 2 , particle *i* update its vector
$\boldsymbol{Vi} = \{(5,2), (9,4), (7,3), (6,5)\}$

PSO Clustering Algorithm

The PSO Clustering Algorithm [4] **pseudocode as follow:**

Initialize each particle with ***K*** random cluster centers.
for iteration count = 1 to maximum iterations **do**
for all particle *i* **do**
for all pattern **X***p* in the dataset **do**
calculate Euclidean distance of **X***p* with all cluster center
assign **X***p* to the cluster that have nearest center to **X***p*
end for
calculate the fitness function *f*.

$$f = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2$$

end for
Find the personal best and global best position of each particle.
Update cluster center according to velocity and coordinate updating formula of PSO.
end for

K-means Clustering

- Partitioning clustering approach
- Each cluster is associated with a centroid (center point)
- Each point is assigned to the cluster with the closest centroid
- Number of clusters, K, must be specified
- The basic algorithm is very simple [3]

1: Select K points as the initial centroids.
2: **repeat**
3:     Form K clusters by assigning all points to the closest centroid.
4:     Recompute the centroid of each cluster.
5: **until** The centroids don't change

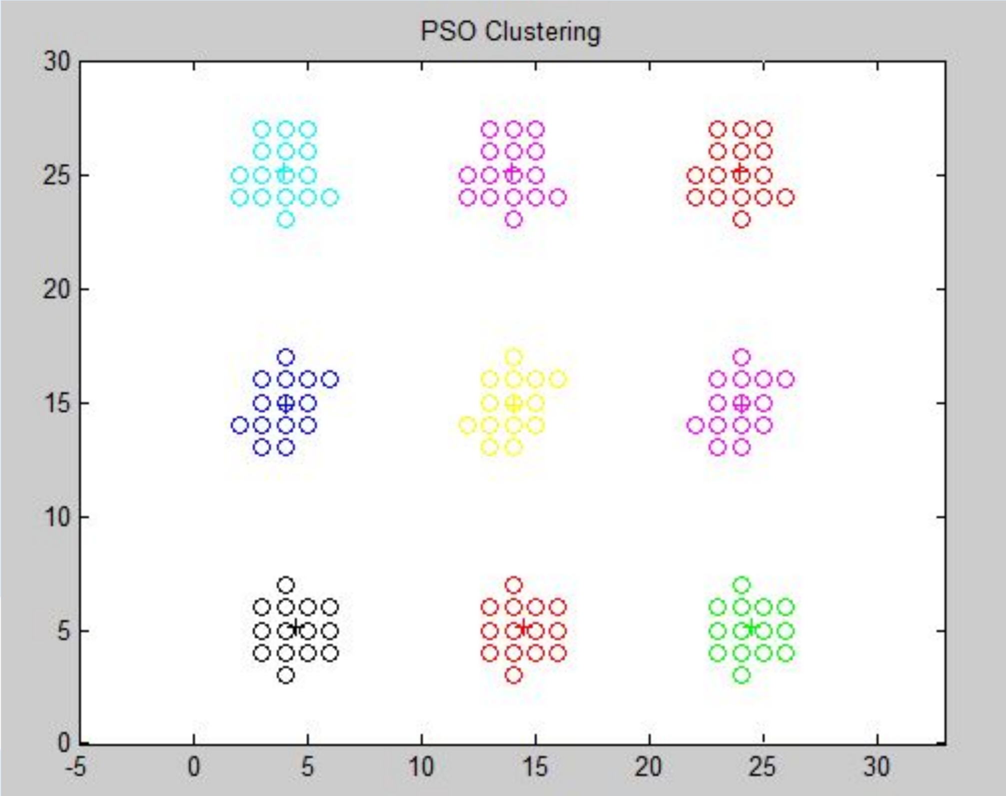
Experimental Results

- The software implemented using Matlab
- PSO clustering algorithm and Kmeans were tested using three type of data set

  - Large data set
  - Small data set
  -Small data set with noisy and outliers

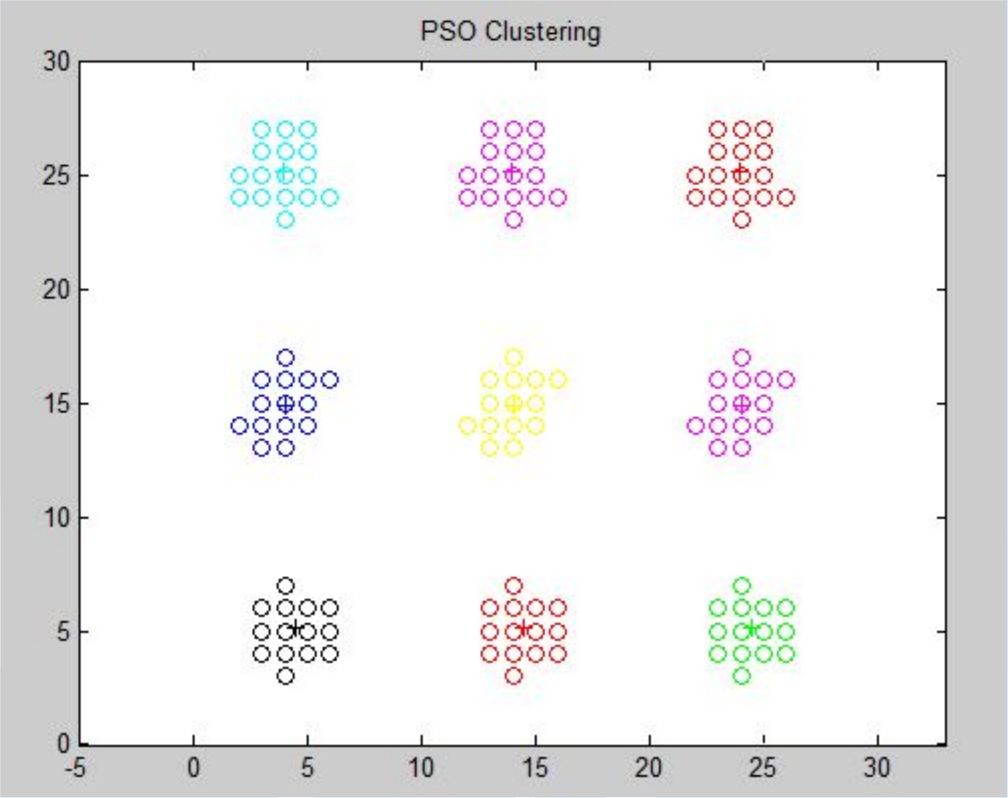
Experimental Results

PSO generate high quality clustering

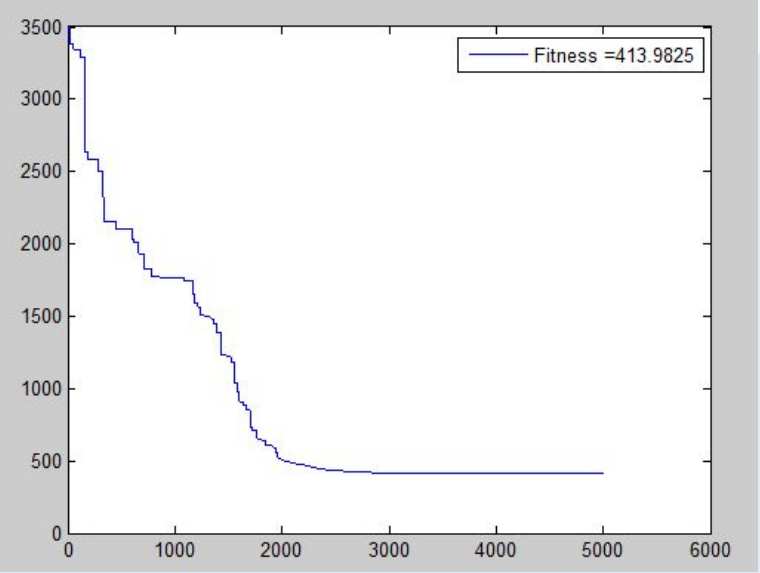


Experimental Results

PSO generate high quality clustering

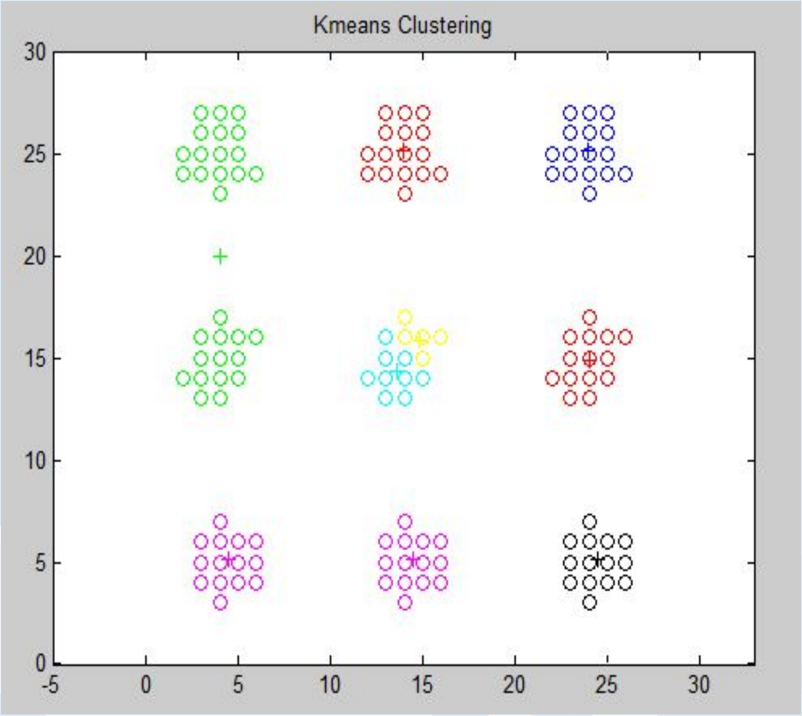


$$\min \ f = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2$$

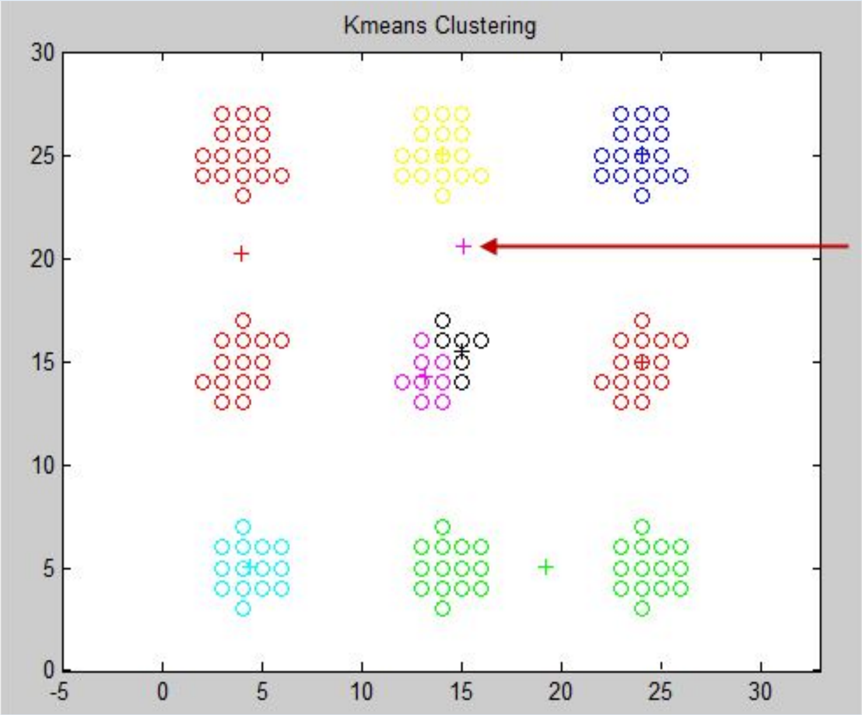


PSO fitness at each iteration

Experimental Results

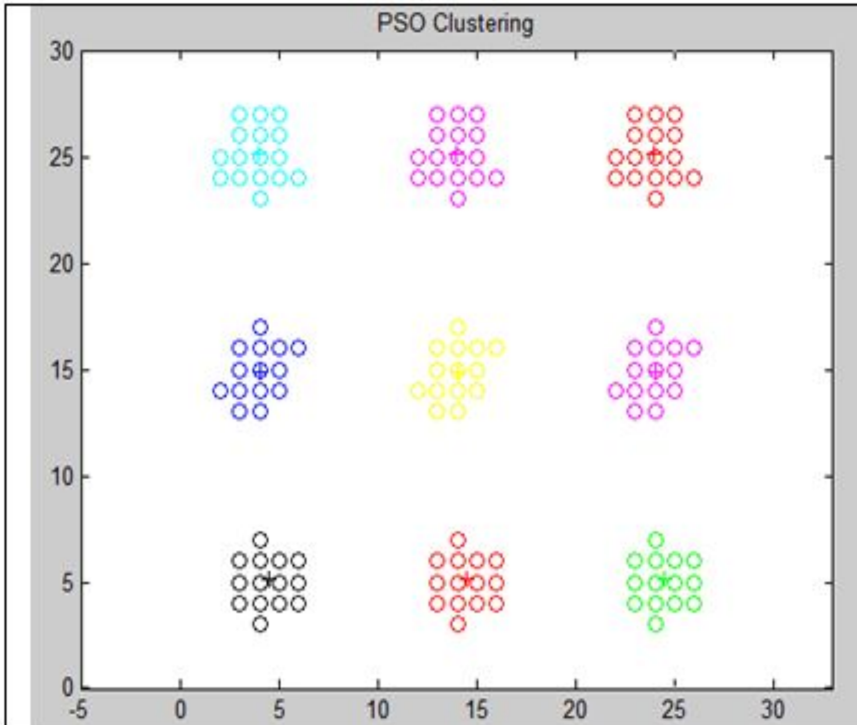


Kmeans terminate at local menimum

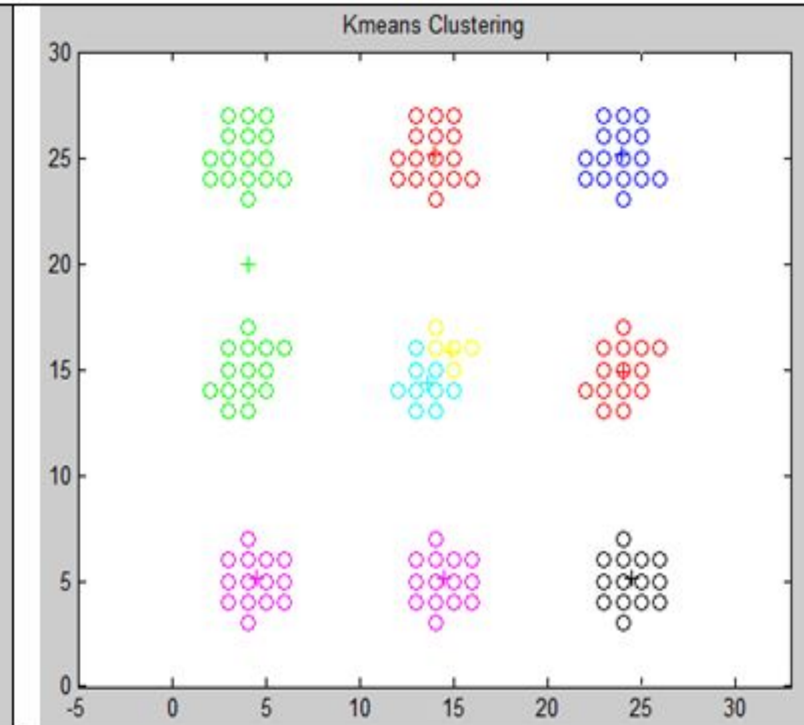


Kmeans generate empty cluster

PSO vs Kmeans



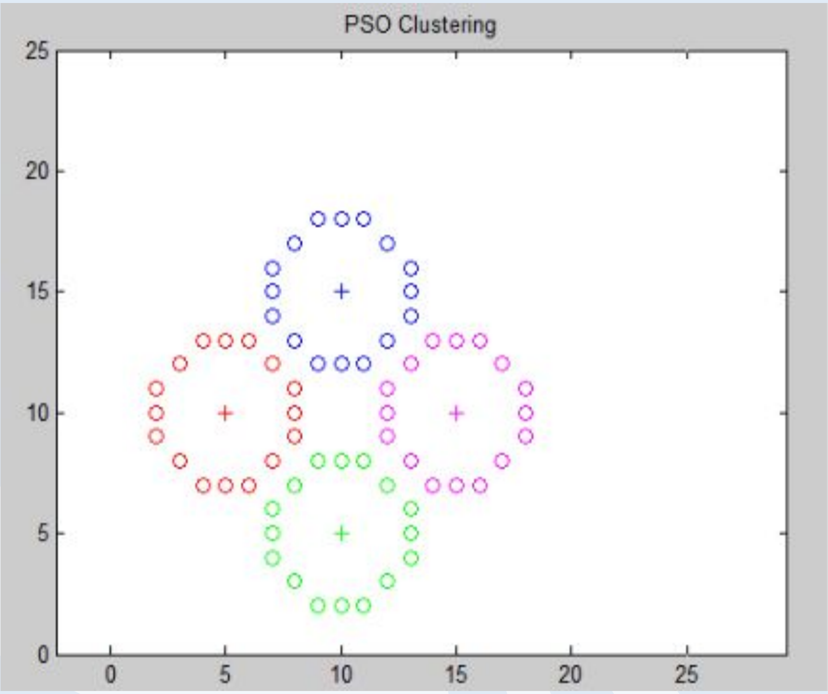
PSO clustering generate 9 clusters



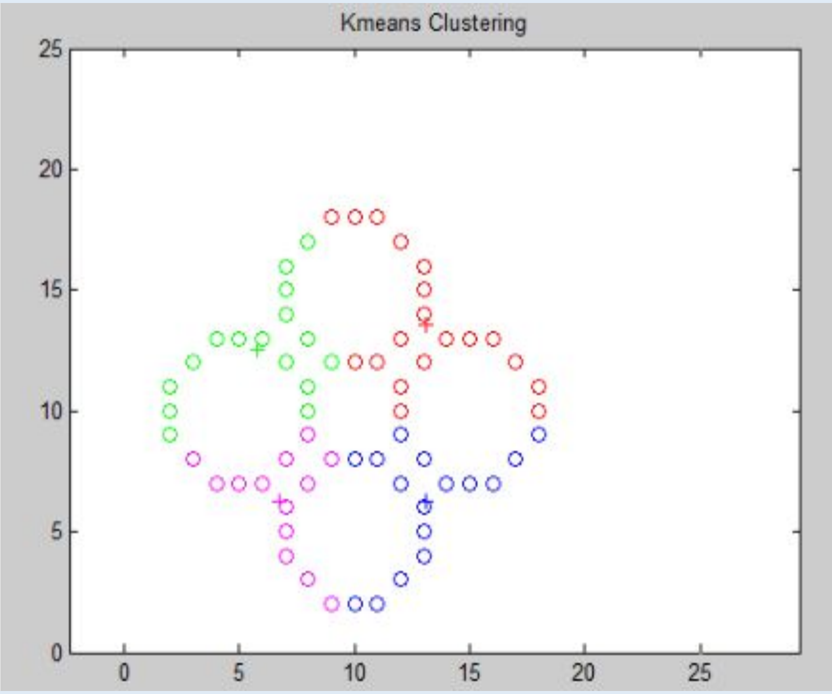
Kmeans clustering generate 9 clusters

PSO vs Kmeans

PSO terminate at global minimum
Kmeans often terminates at local minimum



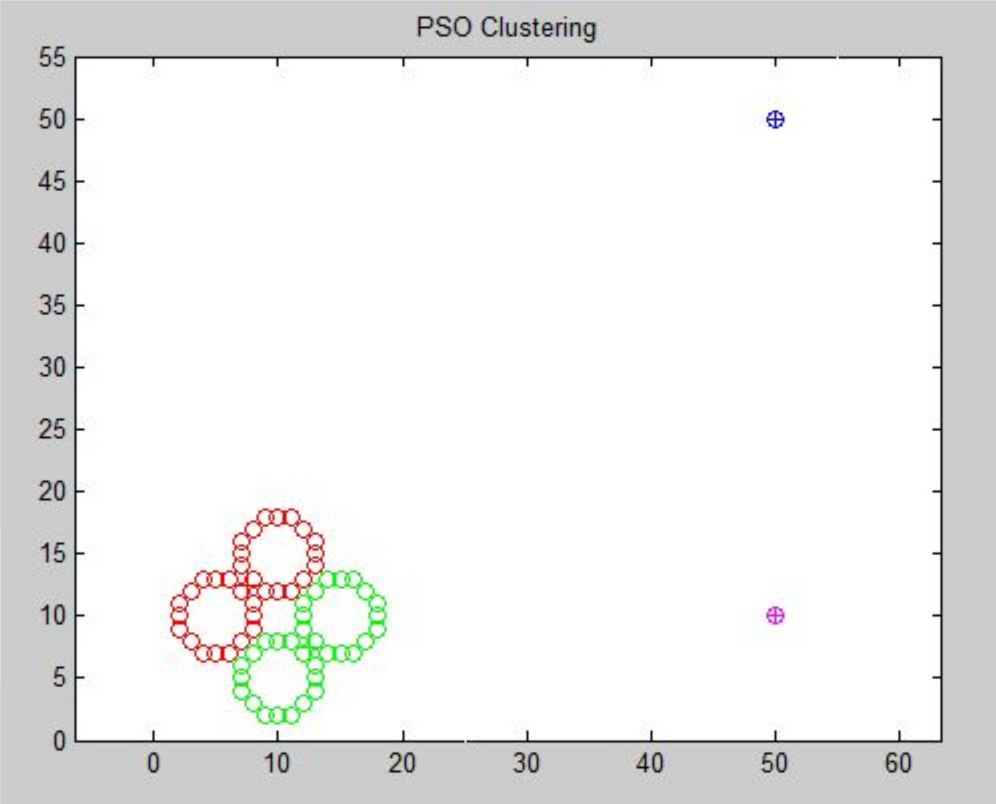
PSO Clustering



Kmeans Clustering

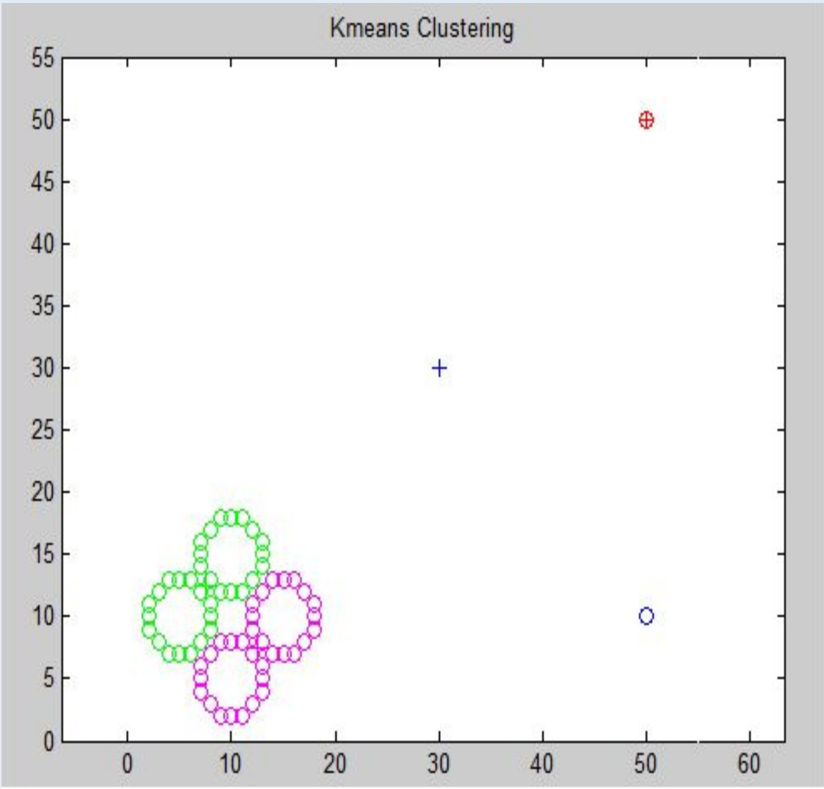
Experimental Results

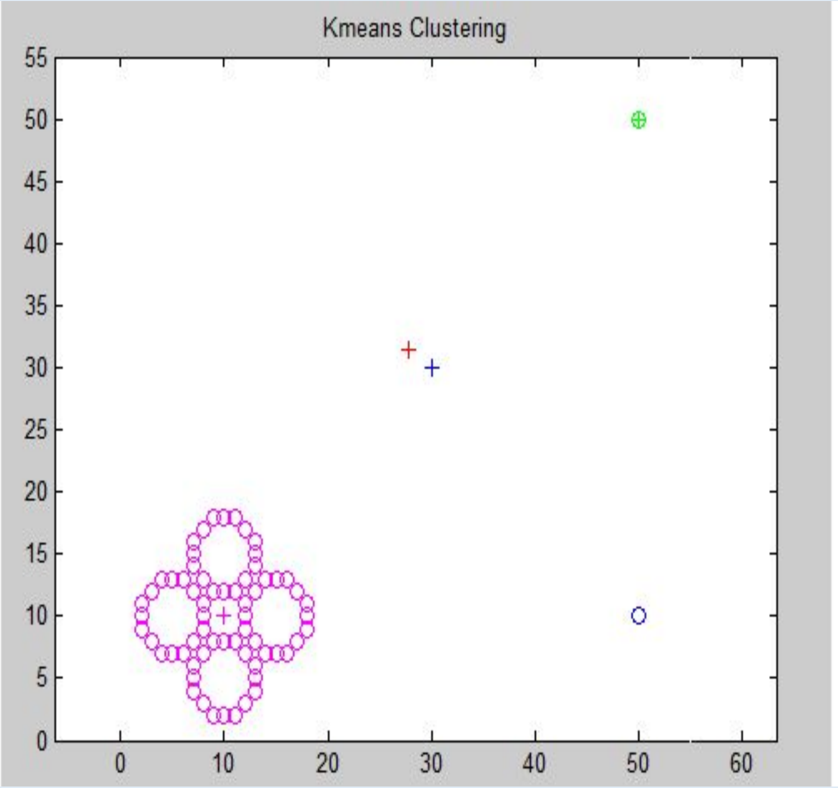
PSO clustering does not affected by noisy and outlier



Kmeans Clustering

Kmeans affected by noisy data and outlier





Conclusion

- **Swarm Intelligence is population based search technique.**
- **PSO is robust stochastic optimization technique and can be applied for data clustering.**
- In **ACO** algorithm, the optimization problem must be written in the form of path finding with a weighted graph
- **PSO clustering algorithm avoid the problems that arise with Kmeans clustering such as terminating at local minimum, generating empty clusters and sensitivity to noisy data and outliers.**

References

[1] Ant Colony Optimization website, http://iridia.ulb.ac.be/~mdorigo/ACO/about.html

[2] J. Kennedy and R.C. Eberhart, “Particle swarm optimization,” in IEEE Int. Conf. on Neural Networks., Perth, Australia, vol. 4, 1995, pp. 1942-1948.

[3] J. Ham and M. Kamber, “Data mining: concepts and techniques (2nd edition,” Morgan Kaufman Publishers, pp. 1-6, 2006.

[4] Van der Merwe, D. W. and Engelbrecht, A. P. “Data clustering using particle swarm optimization”. Proceedings of IEEE Congress on Evolutionary Computation 2003 (CEC 2003), Canbella, Australia. pp. 215-220, 2003.

[5] E. Bonabeau, M. Dorigo, and G. Theraulaz. Swarm Intelligence: From Natural to Artificial System. Oxford University Press, New York, 1999

Questions

Q1: Define Swarm Intelligence and what is the characteristics of the swarm?

Q2: What is the difference between heuristic and metaheuristic?

Q3 What are the types of search techniques and mentioned the components of the search technique?

Q4: What is the Particle Swarm Optimization and show the algorithm?

Q5: Define Ant Colony Optimization and show the algorithm?