Social Spider Optimization
&
Colliding Bodies Optimization
- Application to Data Clustering

Dr. Satyasai Jagannath Nanda, Senior Member IEEE
Dept. of Electronics & Communication Engineering
Malaviya National Institute of Technology Jaipur
Rajesthan-302017, India
Email : sjnanda.ece@mnit.ac.in, +91-9549654237

Data Clustering as an Optimization Problem

Data Clustering refers to the unsupervised classification of a dataset into groups, such that elements within a group are similar to each other.

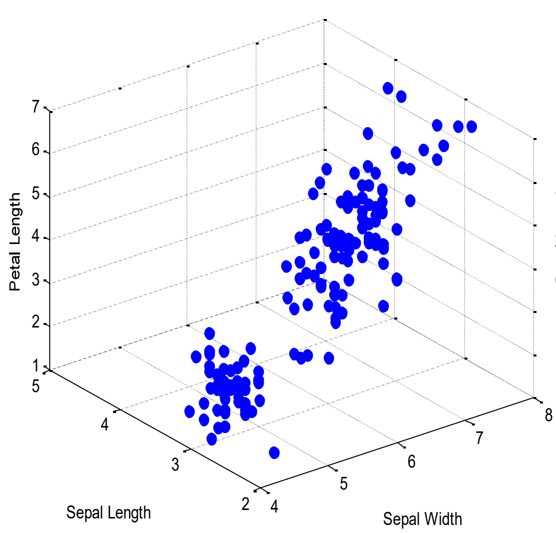
Example : IRIS Flower Data Set

- ✓ 150 Observations (50 Observations for each flower)
- ✓ 4 attributes (Sepal Length, Sepal Width, Petal length, Petal Width)

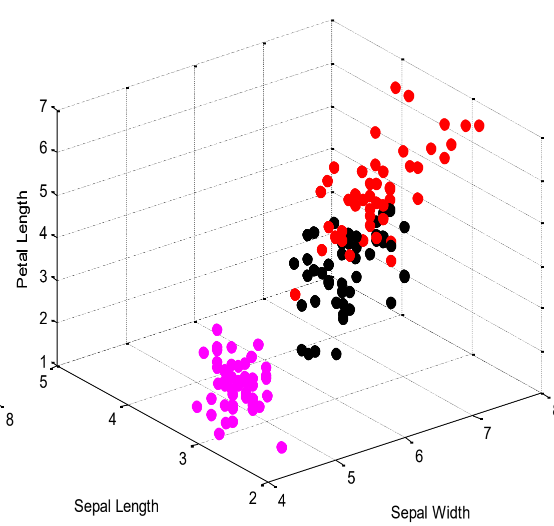

Setosa


Versicolor


Versinica

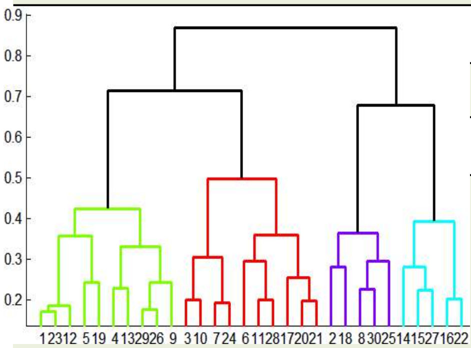
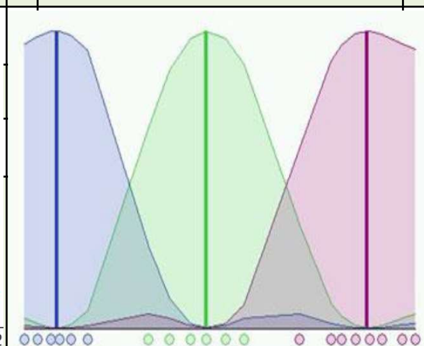
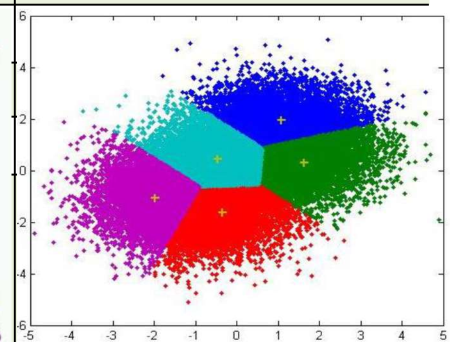


3D Plot of unlabeled IRIS dataset



3D Plot after cluster analysis

Data Clustering

Hierarchical Clustering	Fuzzy Clustering	Partitional Clustering
Dataset is decomposed into several groups and output is expressed in the form of Dendogram plot.	Each element of a dataset belongs to all the groups with a fuzzy membership grade.	Dataset is divided into a number of partitions based upon certain criterion (fitness measure)
		
Popular Algorithms		
✓ Agglomerative algorithms ✓ Divisive algorithms	✓ Fuzzy C-Means (FCM) ✓ Fuzzy C-Shells (FCS)	✓ K-Means ✓ K-Medoids

Data Clustering

Hierarchical Clustering	Fuzzy Clustering	Partitional Clustering
Dataset is decomposed into several groups and output is expressed in the form of Dendogram plot.	Each element of a dataset belongs to all the groups with a fuzzy membership grade.	Dataset is divided into a number of partitions based upon certain criterion (fitness measure)
Effectiveness		
Low Dimensional Dataset	Overlapping Dataset	High Dimensional Datasets
Challenges		
✓ Extensive computation to create dendogram. ✓ Ineffective for overlapping dataset	✓ De-fuzzification process at times introduces more complexity	✓ **Will explain in-details with the progress in this presentation**.
Popular Algorithms		
✓ Agglomerative algorithms ✓ Divisive algorithms	✓ Fuzzy C-Means (FCM) ✓ Fuzzy C-Shells (FCS)	✓ K-Means ✓ K-Medoids

Partitional Clustering

- Given a dataset

$$P_{N\times D} = \begin{bmatrix} p_{1,1} & p_{1,2} & .. & p_{1,D} \\ p_{2,1} & p_{2,2} & .. & p_{2,D} \\ :: & :: & :: & :: \\ p_{N,1} & p_{N,2} & .. & p_{N,D} \end{bmatrix} = \begin{bmatrix} \vec{p_1} \\ \vec{p_2} \\ :: \\ \vec{p_N} \end{bmatrix}$$

- It represents N patterns each having D attributes (also called features)
- **Objective** : Partition the dataset into M groups (Clusters) such that
  - Every cluster must contain at least one pattern

$$C_i \neq \phi \quad \forall i \in \{1,2,..M\}$$

  - There should not be common patterns between any two clusters

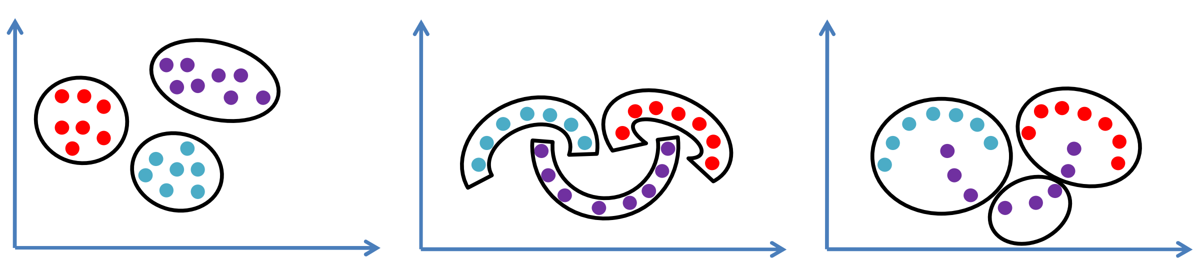
$$C_i \cap C_j = \phi \quad \forall i,j \in \{1,2,..M\} \ \ and \ \ i \neq j$$

  - Dataset consists of sum of all the cluster

$$\bigcup_{i=1}^{M} C_i = P_{N\times D}$$

Partitional Clustering

- **K-means** based Partitional clustering (Loyad-1963, Bell Lab)
  - Uses minimization of **Eucliden distance** as the similarity measure
  - Effective for **circular/spherical shape cluster**
- **Density based** based Partitional clustering (Easter et al. - 1996)
  - Formulate smaller clusters based on **the density of data in a region**
  - Merge the small clusters to formulate large clusters
  - Effective for **arbitrary shape clusters**



Spherical shape cluster **Arbitrary shape cluster** **Problem**

Nature Inspired Algorithms for Clustering

Type	Single Objective			Multi- Objective	
	Algorithm	Inventor	Clustering	Algorithm	Clustering
Evolutionary Algorithms	Genetic Algorithm (GA)	Holland-1975	Bezdek-1994	NSGA-II	Bandyopadhyay-2007
	Evolutionary Strategy (ES)	Rechenberg-1989	Babu & Murty-1994	MO-ES	Costa-2002
	Genetic Programming (GP)	Koza-1992	Falco-2006	MO-GP	Coelho-2010
	Differential Evolution (DE)	Storn-1997	Paterlini-2004	MO-DE	Suresh-2009
Physical Algorithms	Simulated Annealing (SA)	Kirkpatrick-1983	Selim -1991	MO-SA	Bandyopadhyay-2008
	Memetic Algorithm (MA)	Moscato-1989	P. Merz-2003	MO-MA	Knowles-2000
	Harmony Search (HS)	Geem-2001	Mahdavi-2008	MO-HS	Ricart-2011
	Shuffled Frog-Leaping Algorithm (SFL)	Eusuff-2006	Amiri-2009	MO-SFL	J. Liu-2012
Swarm Intelligence	Ant Colony Opt. (ACO)	Dorigo-1992	Shelokar-2004	MO-ACO	Santos-2009
	Particle Swarm Opt. (PSO)	Kennedy-1995	Omran-2002	MO-PSO	S. Das-2008

Nature Inspired Algorithms for Clustering

Type	Single Objective			Multi- Objective	
	Algorithm	Inventor	Clustering	Algorithm	Clustering
Swarm Intelligence	Artificial Bee Colony (ABC)	Basturk-2006	Zhang-2010	MO-ABC	Akbari-2012
	Fish Swarm Algorithm(FSA)	Li et al.-2002	Cheng-2009	MO-FSA	Jiang-2011
Bio-inspired Algorithms	Artificial Immune System(AIS)	Charsto-2002	Nasraoui-03	MO-AIS	Gong-2007
	Bacterial Foraging Opt. (BFO)	Passino-2002	Wan-2011	MO-BFO	Niu-2012
	Krill herd Algorithm	Gandomia-12			
Other Nature Inspired Algorithms	Cat Swarm Opt. (CSO)	S C Chu-2006	Santosa-2009	MO-CSO	Pradhan-2012
	Cuckoo Search Algo.(CSA)	X.S. Yang- 2009	Goel-2011	MO-CSA	X.S. Yang- 2012
	Firefly Algorithm (FA)	X.S. Yang- 2009	Senthilnath-2011	MO-FA	X.S. Yang- 2012
	Invasive Weed Optimization Algo. (IWO)	Mehrabian-2006	Chowdhury-2011	MO-IWO	R.Liu-2012
	Gravitational Search (GS)	Rashedi-2007	Hatamlou-2011	MO-GS	Hassanzadeh-2010
	Bat-inspired Algorithm	X.S. Yang- 2010		MO-BAT	X.S. Yang- 2009

Steps of Benchmark Nature Inspired Algorithms

		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
GA	Initialize Chromos.	⇒	Crossover	⇒	Mutation	⇒	Fitness	⇒	Selection	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
DE	Initialize Particles	⇒	Mutation	⇒	Crossover	⇒	Fitness	⇒	Selection	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
ACO	Initialize Ants	⇒	Fitness	⇒	Update Ph. Intensity	⇒	Drop or Peak	⇒	Short Memory	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
PSO	Initialize Particles	⇒	Vel. update Pos. update	⇒	Compute G_{Bst} & P_{Bst}	⇒	Fitness	⇒	Selection	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
ABC	Initialize Bees	⇒	Compute Emp. Bees	⇒	Greedy Sel. & Fitness	⇒	Onlooker Bees	⇒	Selection	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
AIS	Initialize Im. Cells	⇒	Fitness	⇒	Clone	⇒	Mutation	⇒	Selection	⇒	Cl. O/p	
		⇐ ⇓	⇐	⇐	Next Generation		⇐	⇐	⇐ ⇑			
BFO	Initialize Bacteria	⇒	Chemo-taxis	⇒	Swarming	⇒	Reprodu-ction	⇒	Eliminat. & Dispers.	⇒	Cl. O/p	

Fitness Evaluation/Cluster Validity Index

- Cluster validity indices represent statistical functions used for quantitative evaluation of the clusters derived from a dataset (can also can be used as fitness functions)

Statistical Function	Nature for Clustering	Used in Research Papers
Medoid Distance	Minimization	Lucasius -1993, Sheng & Liu-2004
Centroid Distance	Minimization	Murty-1996, Maulik -2000, Zhang-2011
Distortion Distance	Minimization	Krishna & Murty-1999, Kivijarvi-2003, Lu-2004
Variance Ratio Criterion	Maximization	Cowgill-1999, Casillas-2003
Intra and Inter Cluster Distance	Min/Max depend on ratio	Tseng & Yang- 2001, Das- 2008, Nanda-2012 Handl and Knowles- 2004, 2007, 2010
Dunn's index	Maximization	Dunn-1974, Zhang and Cao-2011
Davis-Bouldin index	Minimization	Davis & Bouldin-1974, Bandyopadhyay-2002
CS Measure	Minimization	Chou-2004, Das-2008
Silhouette Index	Maximization	Kaufman-2008, Hruschka-2009

Popular Datasets used for Validation

✓ UCI Machine Learning Database Online - http://archive.ics.uci.edu/ml/datasets.html

Data Set/ Creator	Dimension/ Num of Cl.	Used in Research Papers
Iris R.A. Fisher	150 x 4 Cluster-3	**GA** : Bezdek-1994, R.Xu-2005, Xiao-2010, He-2012, **ACO** : Chen-2005, Handl-2004, Zhang-2011, Wan-2012, **PSO** : Tsai-2011, Yang-2009, Cura-2012, Kuo-2012, **DE** : Das-2008, Hu- 2010, Kwedlo-2011, **CSO** : Santosa-2009 **ABC** : Zhang-2010, Zou-2010, Karaboga-2011, **MOAIS** : Gong-2007
Wine Forina	178 x 13 Cluster-3	**PSO** : Niknam-2010, Chuang-2011, Tsai- 2011, Cura-2012 , **ACO** : Azzag-2003, Ghosh-2008, Wan-2012, Chowdhury-2012, **BFO** : Olesen-2009, **Frog** : Amiri - 2009, **NSGA II** : Ripon-2009,
Glass B. German	214 x 9 Cluster-6	**GA** : Abraham-2007, Xiao- 2010, He-2012, **DE** : Das-2008, Hu-2010, **PSO** : Jarboui-2007, Yang-2009, Tsai-2011, Kuo-2012, **Firefly** : Senthilnath-2011, **Gravitational Search** - Yin-2011
Image Segm. Vision Group	2310 x 19 Cluster-7	**GA** : Hong-2008, **ACO** : Chowdhury-2012, Ghosh- 2008, **DE** : Kwedlo-2011, **ABC** : Zou-2010, Karaboga-2011,
Indian Diabetes V. Sigillito	768 x 8 Cluster- 2	**ACO** : Azzag- 2003, **BFO** : Olesen-2009, **ABC :** Karaboga-2011, **Firefly :** Senthilnath-2011, **NSGA II** : Ripon-2009,

Real Life Applications of Clustering

Applications	Key Research Papers
Image Segmentation	**PSO** : Lee-2012, Abraham-2007, Zhang-2011, **Firefly :** Hassanzadeh-2011, **ACO** : Ghosh-2009, **DE** : Das-2008, **SFL** : Bhaduri-2009, **NSGA II** : Mukhopadhyay-2011
Image clustering	**GA :** Bandyopadhyay-2002, **DE** : Das-2008, Omran-2005, **PSO** : Omran-2002, **NSGA II** : Bandyopadhyay-2007
Document clustering	**GA** : Casillas-2003, Kuo-2010, **PSO** : Cui-2005, **ACO** : Handl-2002, **DE** : Abraham-2006, **HS** : Mahdavi-2009, **Review** : AK Jain-2010, Steinbach-2000
Web mining	**ACO** : Abraham-2003, **PSO** : Alam-2008, **HS** : Mahdavi-2008, **SFL** : Fang- 2011
Text mining	**ACO :** Handl-2002, Vizine-2005,
Wireless Sensor Network	**GA** : Tan-2012, **PSO :** Yu-2011, **ABC** : Udgata-2009, **HS :** Hoang-2010, **MOCSO :** PMPradhan-2012
Mobile Network	**PSO** : Ji -2004, **DE** : Chakraborty-2011, **SA :** Wjin-2005,
Gene expression Data Analysis	**GA :** Lu-2004, Ma-2006, **DE** : Das-2008, **PSO** : Sun-2012, Du-2011, Thangavel-2011, **AIS :** Liu-2012, **MODE :** Suresh-2009, **Review** : Xu-2005, Hruschka-2009, Jain-2010
Intrusion Det.	**GA** : Liu-2004, **ACO :** Tsang -2006, **PSO :** Lima-2010
Geophysics	**Review** - AK Jain-2010, N. Cho-2011

Swarm and Evolutionary Computation 16 (2014) 1–18

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

Review

A survey on nature inspired metaheuristic algorithms for partitional clustering

Satyasai Jagannath Nanda[a,*], Ganapati Panda[b]

[a] Department of Electronics and Communication Engineering, Malaviya National Institute of Technology Jaipur, Rajasthan 302017, India
[b] School of Electrical Sciences, Indian Institute of Technology Bhubaneswar, Odisha 751013, India

A R T I C L E I N F O

Article history:
Received 10 October 2012
Received in revised form
23 August 2013
Accepted 20 November 2013
Available online 18 January 2014

Keywords:
Partitional clustering

A B S T R A C T

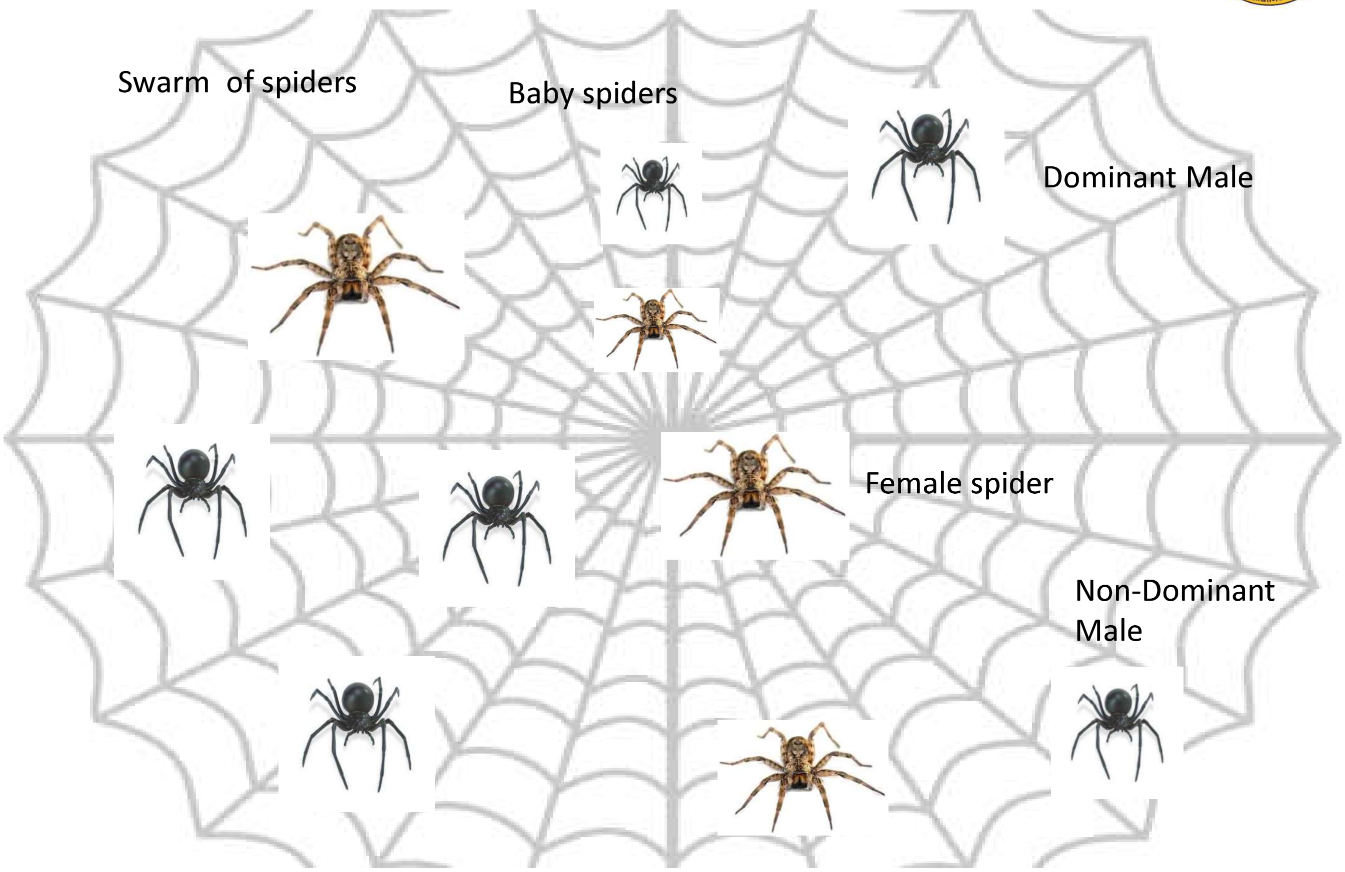
The partitional clustering concept started with K-means algorithm which was published in 1957. Since then many classical partitional clustering algorithms have been reported based on gradient descent approach. The 1990 kick started a new era in cluster analysis with the application of nature inspired metaheuristics. After initial formulation nearly two decades have passed and researchers have developed numerous new algorithms in this field. This paper embodies an up-to-date review of all major nature inspired metaheuristic algorithms employed till date for partitional clustering. Further, key issues involved during formulation of various metaheuristics as a clustering problem and major application areas are discussed.

Social Spider Algorithm

- Swarm based approach
- Search space : Communal Web
- Two different agents: Male and Female spider
- Female biased populations
- Solution : position of the spider
- Fitness : Weight of spider
- Communication : Web

1) E. Cuevas, M. Cienfuegos, D. Zaldívar and M. Pérez-Cisneros, A swarm optimization algorithm inspired in the behavior of the social spider. Expert Systems with Applications, Elsevier, vol. 40, no. 16, pp. 6374-6384, 2013.
2) E. Cuevas, M. Cienfuegos, R. Rojas, A. Padilla, “A computational intelligence optimization algorithm based on the behavior of the social-spider”. Computational Intelligence Applications in Modeling and Control, Springer, pp. 123–146, 2015.

Social Spider Algorithm



Social Spider Algorithm

1. **60–90%** of the total spider population is composed of **females**.
2. The **swarm size is constant**.
3. **Fitness** of each spider is imitated in terms of their **weight**.
4. The **communication is designed with help of vibration on the web**.
5. The Percentage Factor (PF) determines the **female socialization behavior and it is a constant with a value of 0.7 (mostly taken).**
6. Mating is carried out when **a dominant male is successfully able to locate a group of female in his mating radius** within search space.

Social Spider Algorithm

1. Population Initialization

- P= Total number of spiders
- N_F = floor(0.9 - rand x (0.25)) x P
- N_M=P-N_F N_F= no of females N_M= no of males
- T is the dimensionality $S_{P\times T} = [S_{M\times T} \bigcup S_{F\times T}]$

2. Fitness Evaluation (Intra Cluster Distance)

$$C_K = min \sum_{i=1}^{K} \delta[(R_{j,D}, S_{1+(D\times(i-1)),D\times i})]^2; \quad \forall j \in [1, N]$$

Minimization of Euclidean Distance between each data point to cluster center, where the spider position represents the cluster nodes

Social Spider Algorithm

3. Assigning Weight to the Spiders

$$W(i) = \frac{worst_f - f(i)}{worst_f - best_f}$$

Maximum weight is the best spider in the colony

4. Male population fragmentation

$$S_M(i) = \begin{cases} S_{DM}(i) & \text{if } W(i) \geq median(W), \\ S_{NM}(i) & \text{otherwise.} \end{cases}$$

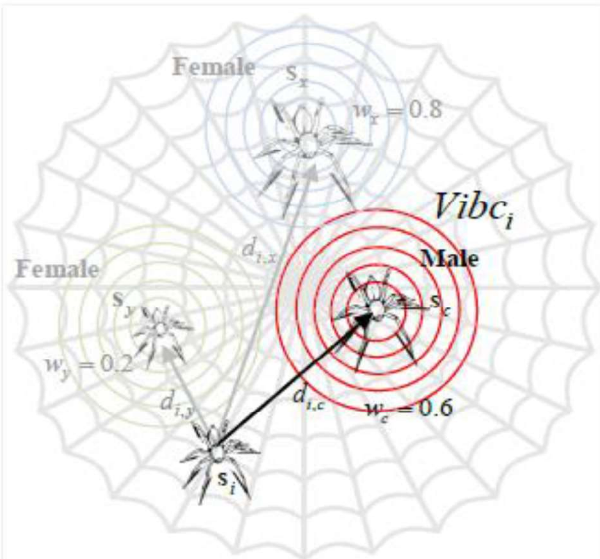
Communication Among Spiders

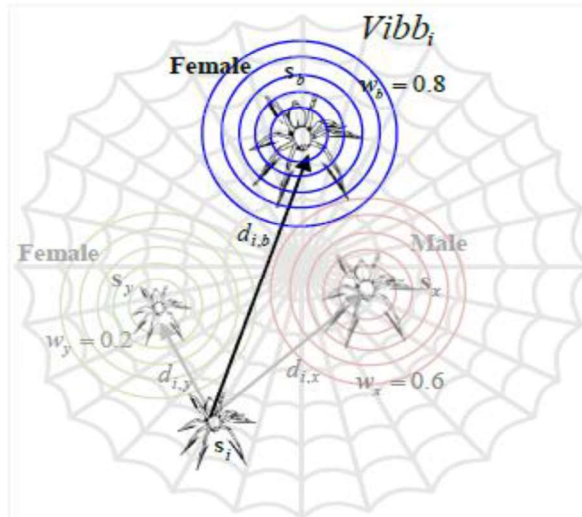
Vibrations received by ith spider from the nearest (or **local spider**)

$$\vartheta_{i,l} = W_l e^{-\delta(S(i),S^l)^2}$$

Vibrations received by ith spider from the spider having highest weight among all (or **global spider**)

$$\vartheta_{i,g} = W_g e^{-\delta(S(i),S^g)^2}$$

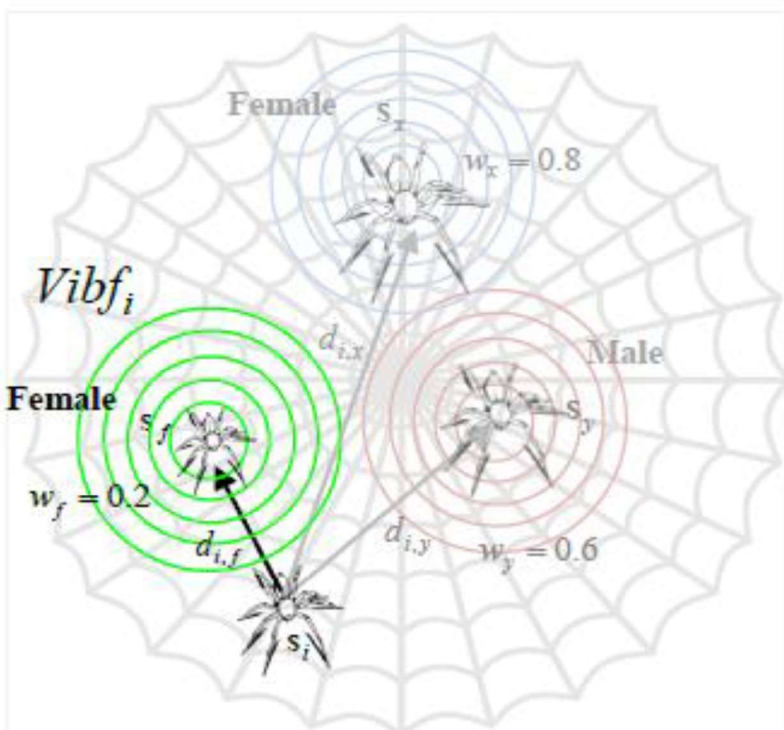




Communication Among Spiders

Vibrations received by ith dominant male spider from the nearest **female spider**

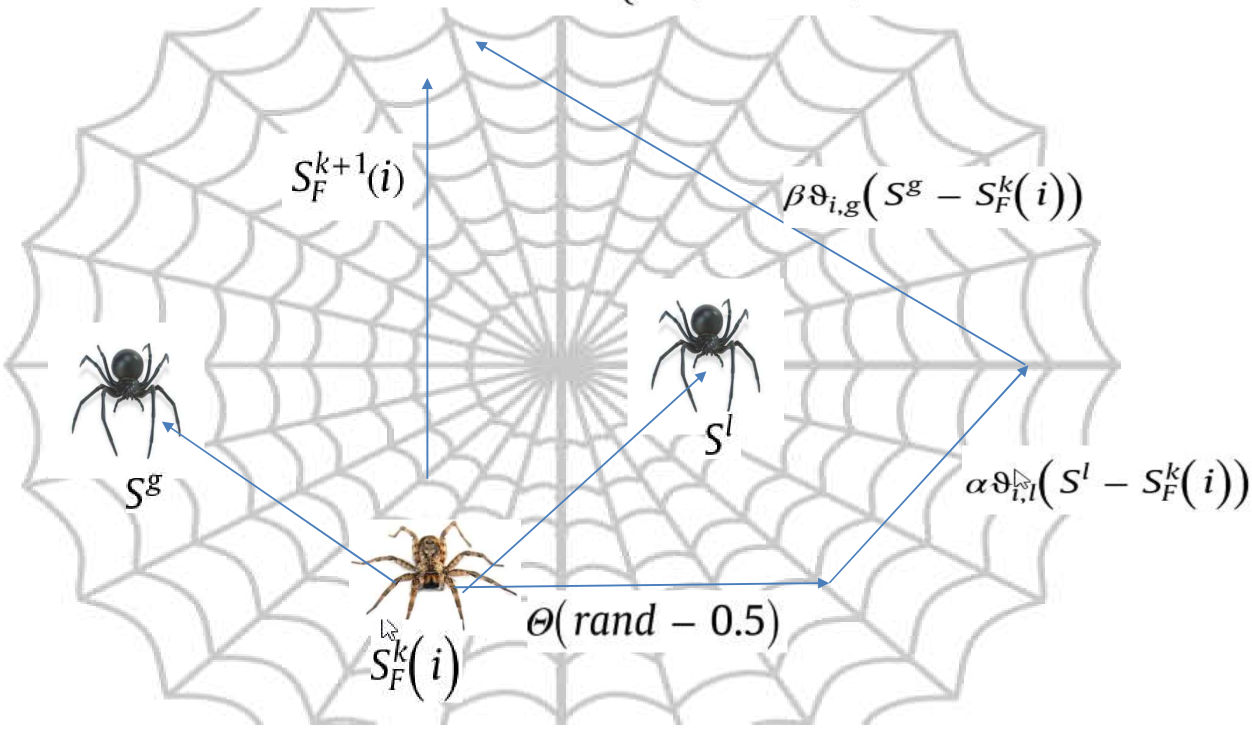
$$\vartheta_{i,F} = W_F e^{-\delta(S(i),S_F)^2}$$



5.Female Moments

Exploration

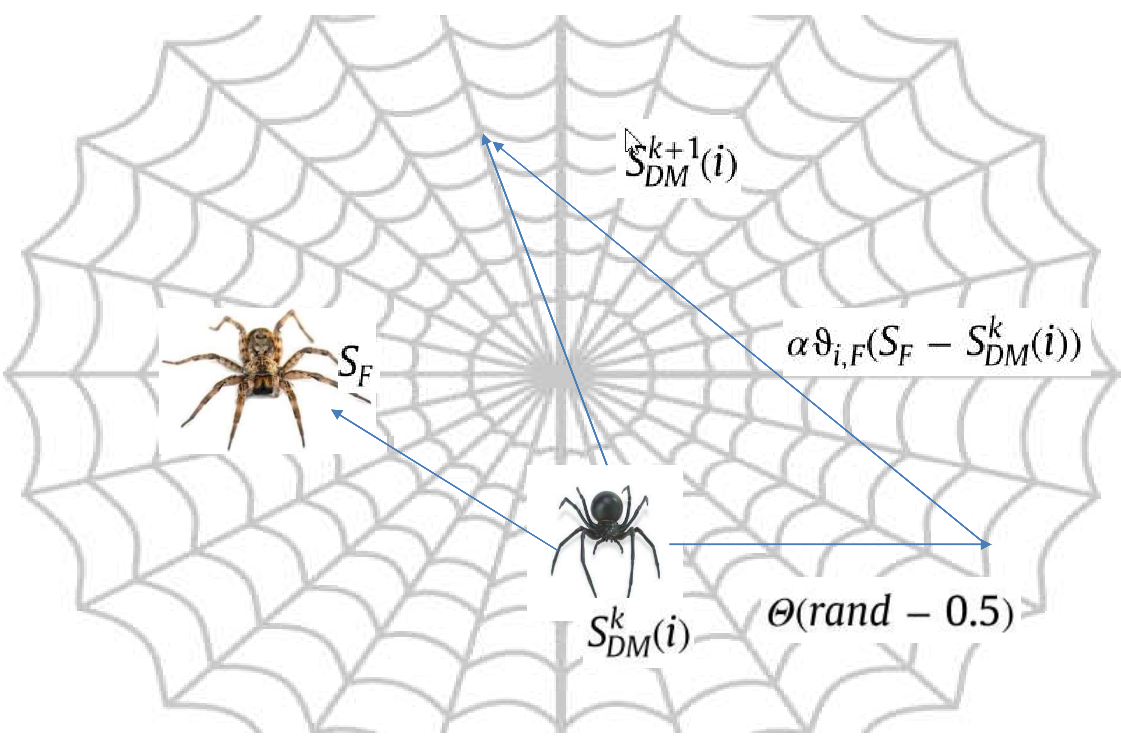
$$S_F^{k+1}(i) = \begin{cases} S_F^k(i) + \alpha\vartheta_{i,l}\left(S^l - S_F^k(i)\right) + \beta\vartheta_{i,g}\left(S^g - S_F^k(i)\right) + \dots \\ +\Theta(rand - 0.5) \quad \text{if } rand < \text{PF} \\ S_F^k(i) - \alpha\vartheta_{i,l}\left(S^l - S_F^k(i)\right) - \beta\vartheta_{i,g}\left(S^g - S_F^k(i)\right) + \dots \\ +\Theta(rand - 0.5) \quad \text{otherwise} \end{cases}$$



6. Dominant Male Movements

Mating

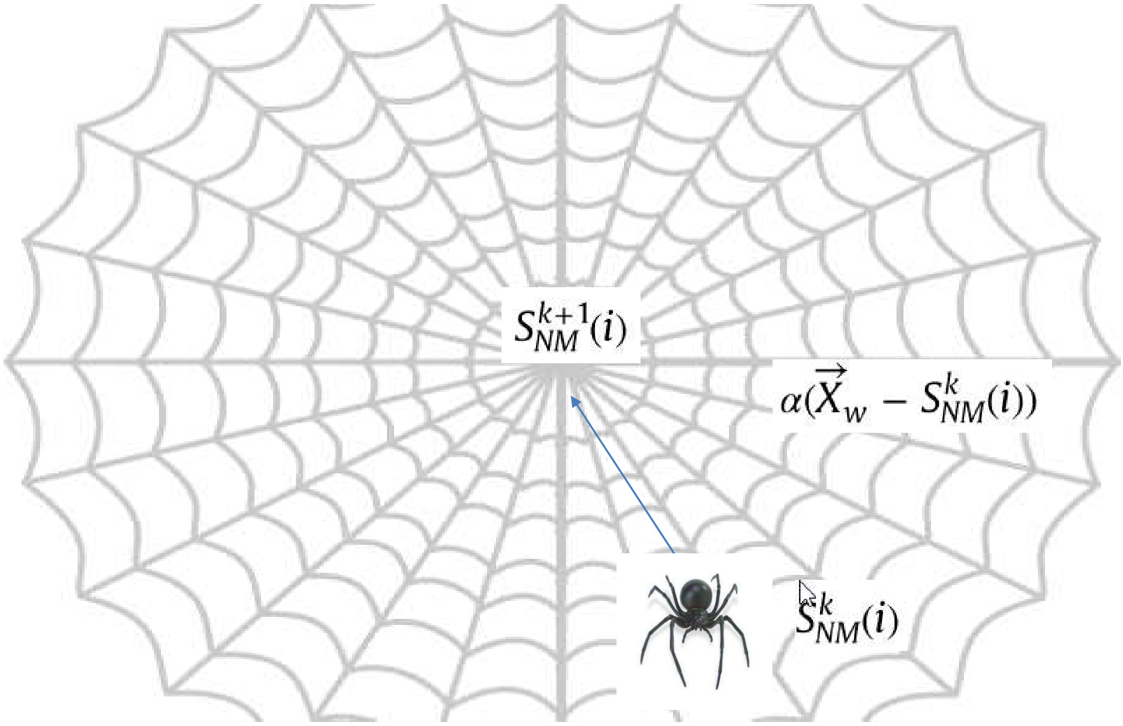
$$S_{DM}^{k+1}(i) = S_{DM}^k(i) + \alpha\vartheta_{i,F}(S_F - S_{DM}^k(i)) + \Theta(rand - 0.5)$$



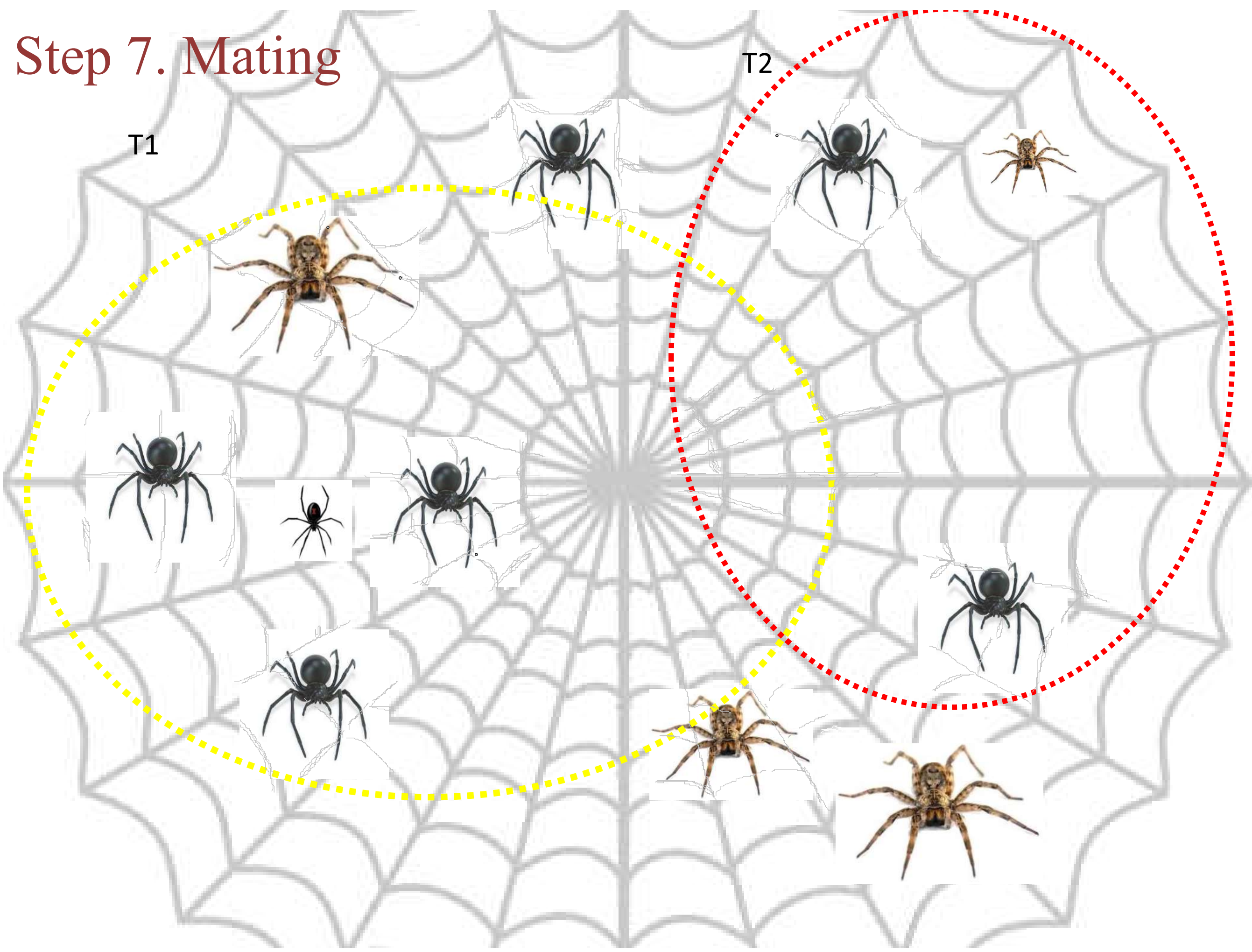
6. Non-Dominant Male Movements

Exploitation

$$S_{NM}^{k+1}(i) = S_{NM}^{k}(i) + \alpha(\overrightarrow{X}_w - S_{NM}^{k}(i))$$

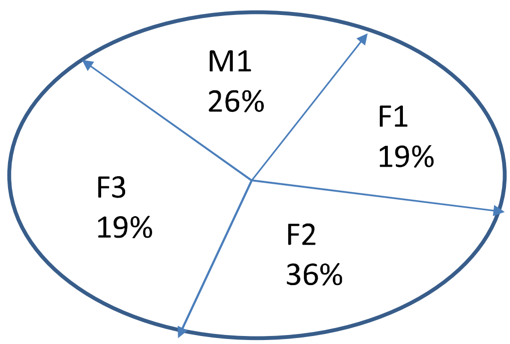


Step 7. Mating

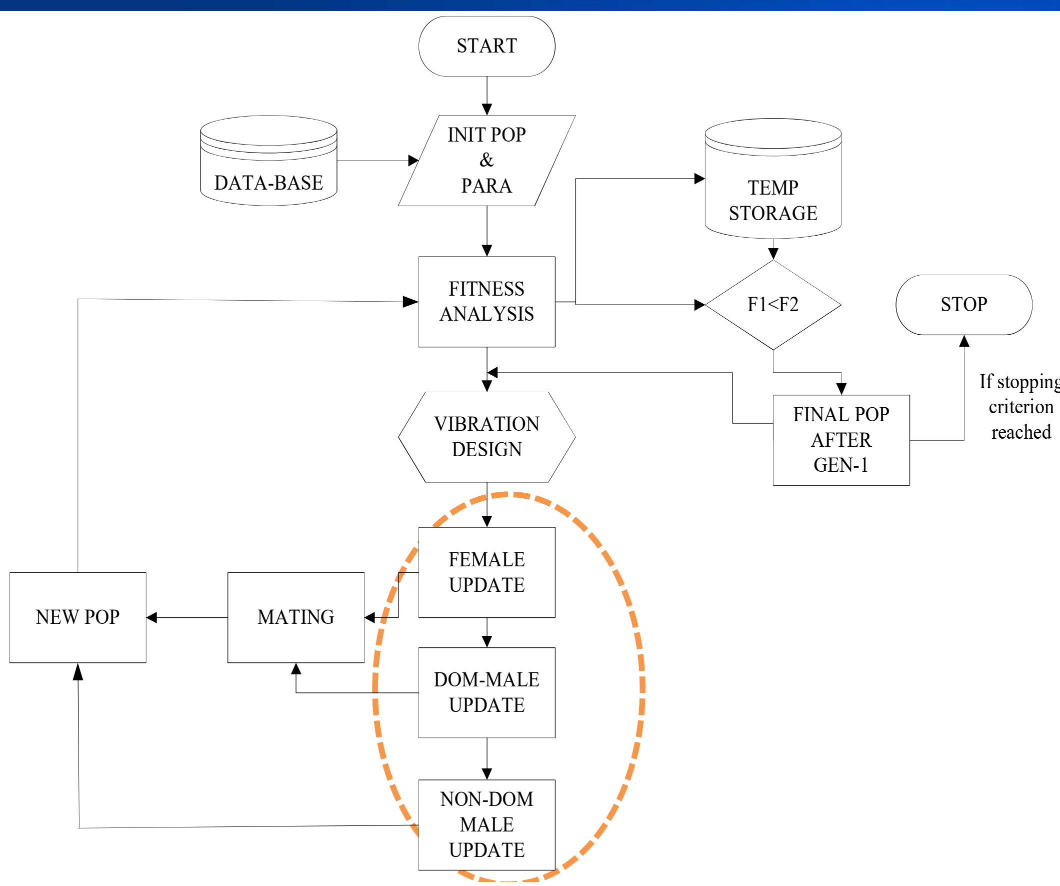


8. Mating for T1 set

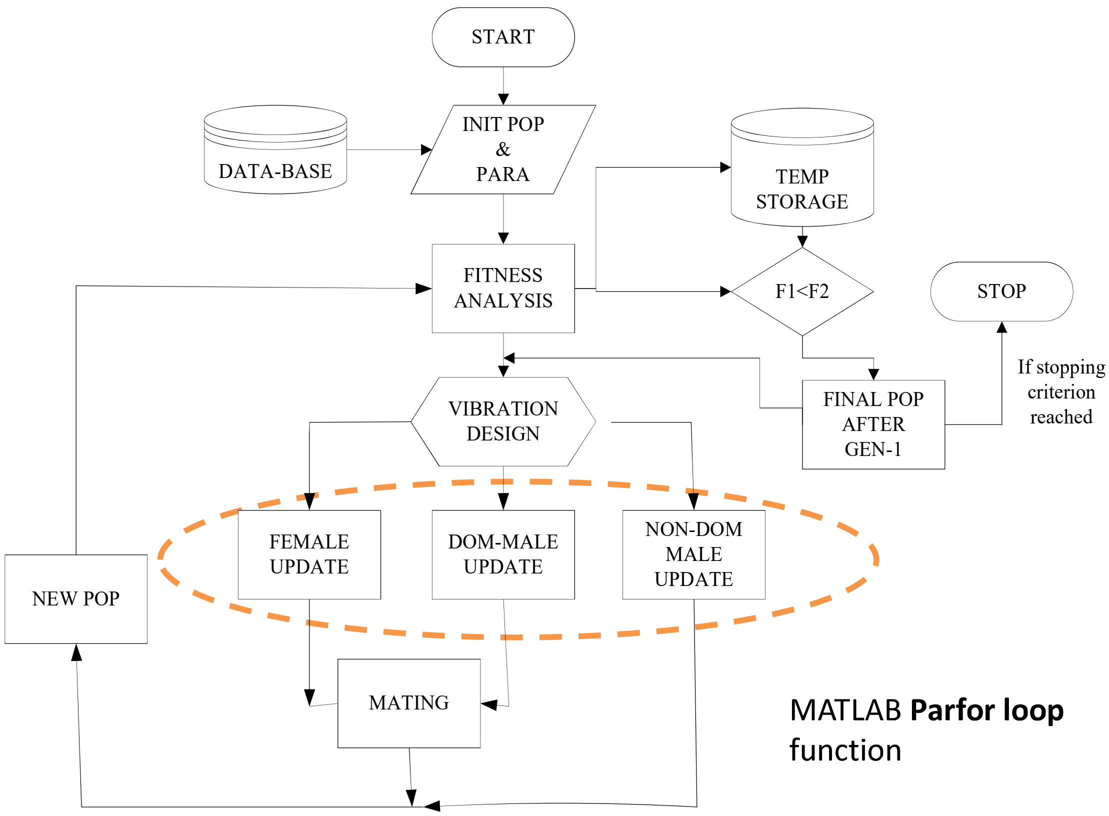
Spider		Position	Weight	P(si)
S1	M1	0.2	0.4	0.26
S2	F1	0.45	0.3	0.19
S3	F2	0.1	0.5	0.36
S4	F3	0.02	0.29	0.19
			1.490	
Snew		0.1	0.5	



Original Social Spider Algorithm



Parallel Social Spider Algorithm



Benchmark Datasets Used for Analysis

Small Dimensional

Sr.no	Data Set	Observation	Attributes	Class	Train	Test
1	Iris	150	4	3	112	38
2	Balance	625	4	3	469	156
3	Wine	178	13	3	133	45
4	Breast Cancer	699	10	2	524	175
5	Thyroid	215	5	3	162	53
6	Credit	690	14	2	518	172
7	Heart	303	14	5	227	76
8	Glass	214	10	7	161	53
9	Horse	364	28	3	273	91

Labeled High Dimensional

Sr.no	Data Set	Observation	Attributes	Class
1	DIM1	1024	1024	16
2	DIM2	1024	32	16
3	DIM3	1024	512	16
4	KDD CUP04 BIO	145741	74	10

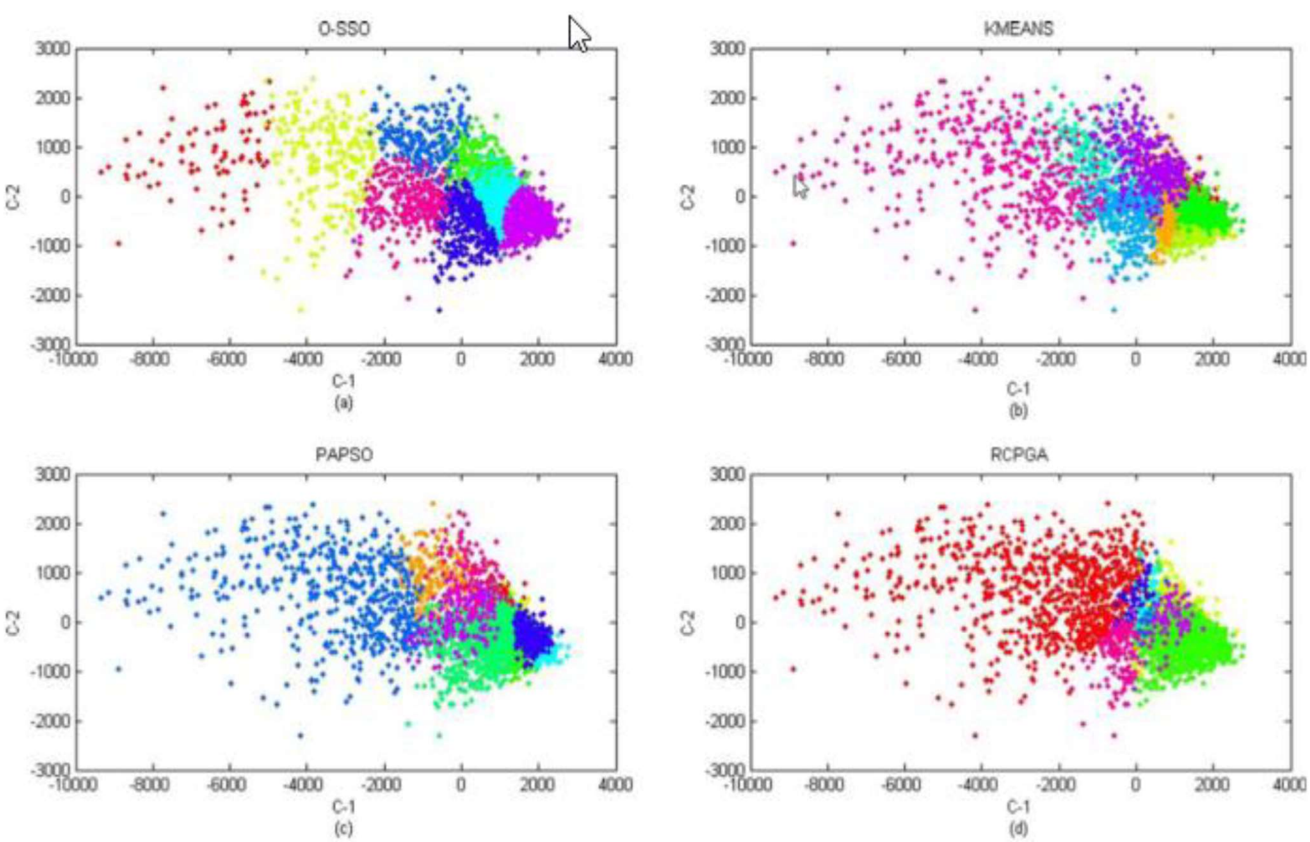
Results Analysis

No of iteration :1000
Independent runs: 10
Evaluation Metric :CEP (Classification error percentage) Rate of miss-classification

Sr.no	Data Set	SSO	ABC [1]	PSO [2]	RGA	k-mean
1	Iris	**0**	0	2.63	0	0
2	Balance	**30.12**	15.38	25.47	28.20	32.69
3	Wine	**0**	0	2.22	22.72	15.90
4	Breast Cancer	**1.32**	0	2.87	25.24	41.37
5	Thyroid	**3.77**	3.77	5.55	18.86	5.66
6	Credit	**21.25**	13.37	22.96	20.26	30.81
7	Heart	**15.66**	14.47	17.46	55.21	56.00
8	Glass	**32.07**	41.50	39.05	52.63	50.94
9	Horse	**36**	38.26	40.98	30.12	29.33

[1] D. Karaboga, C. Ozturk, A novel clustering approach: Artificial bee colony (abc) algorithm, Applied soft computing 11 (1) (2011) 652–657.
[2] I. De Falco, A. Della Cioppa, E. Tarantino, Facing classification problems with particle swarm optimization, Applied Soft Computing 7 (3) (2007) 652–658.

KDD CUP04 BIO Dataset with 10 Clusters



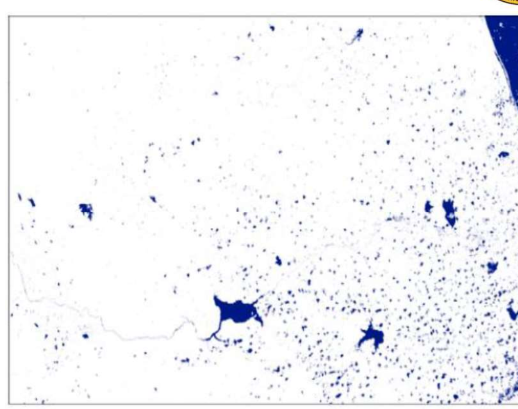
Computational Time Analysis

Sr.no	Parameter Setting	Value
1	Pop size	50
2	No of independent iteration	10
3	Mating probability	75%

		P-SSO		O-SSO	
Sr.no	Dataset	Time (sec)	CEP	Times (sec)	CEP
1	Statlog	**25.0623**	7.3055	44.4226	8.455
2	Handwritten Digit	**25.0435**	2.7988	35.3781	3.5313
3	Dermatology	**1.4142**	13.3880	1.626	14.235
4	Mice	**5.0680**	6.3889	7.6621	13.250

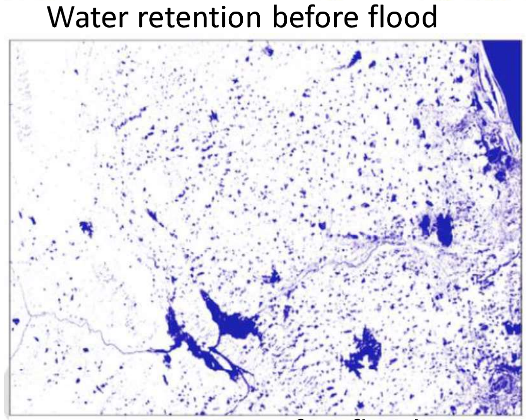
Ex. Analysis on Flood Affected Area of Chennai

Image Captured on Oct. 21, 2015



Water retention before flood

Image Captured on Dec. 08, 2015



Water retention after flood

Publication & References

1) U. P. Shukla and S. J. Nanda, Parallel social spider clustering algorithm for high dimensional datasets. **Engineering Applications of Artificial Intelligence**; **Elsevier**, vol. 56, pp : 75-90, 2016.
2) U. P. Shukla and S. J. Nanda, A Binary Social Spider Optimization Algorithm for Unsupervised Band Selection in Compressed Hyperspectral Images, **Expert Systems with Applications, Elsevier**, vol. 97, pp: 336-356, 2018.
3) U. P. Shukla and S. J. Nanda. "Dynamic clustering with binary social spider algorithm for streaming dataset." **Soft Computing , Springer,** vol. 23, no. 21 , pp : 10717-10737, 2019.
4) U. P. Shukla and S. J. Nanda. "Designing of a Risk Assessment Model for Issuing Credit Card Using Parallel Social Spider Algorithm." **Applied Artificial Intelligence, Taylor & Francis***,* vol. 33, no. 3, pp: 191-207, 2019.
5) R. Gupta, S. J. Nanda, and U. P. Shukla. "Cloud detection in satellite images using multi-objective social spider optimization." **Applied Soft Computing, Elsevier,** vol. 79 pp : 203-226, 2019.
6) S. Aggarwal, P. Chatterjee, R. P. Bhagat, K. K. Purbey, S. J. Nanda, A Social Spider Optimization Algorithm with Chaotic Initialization for Robust Clustering, **Procedia Computer Science, Elsevier**, 2019.

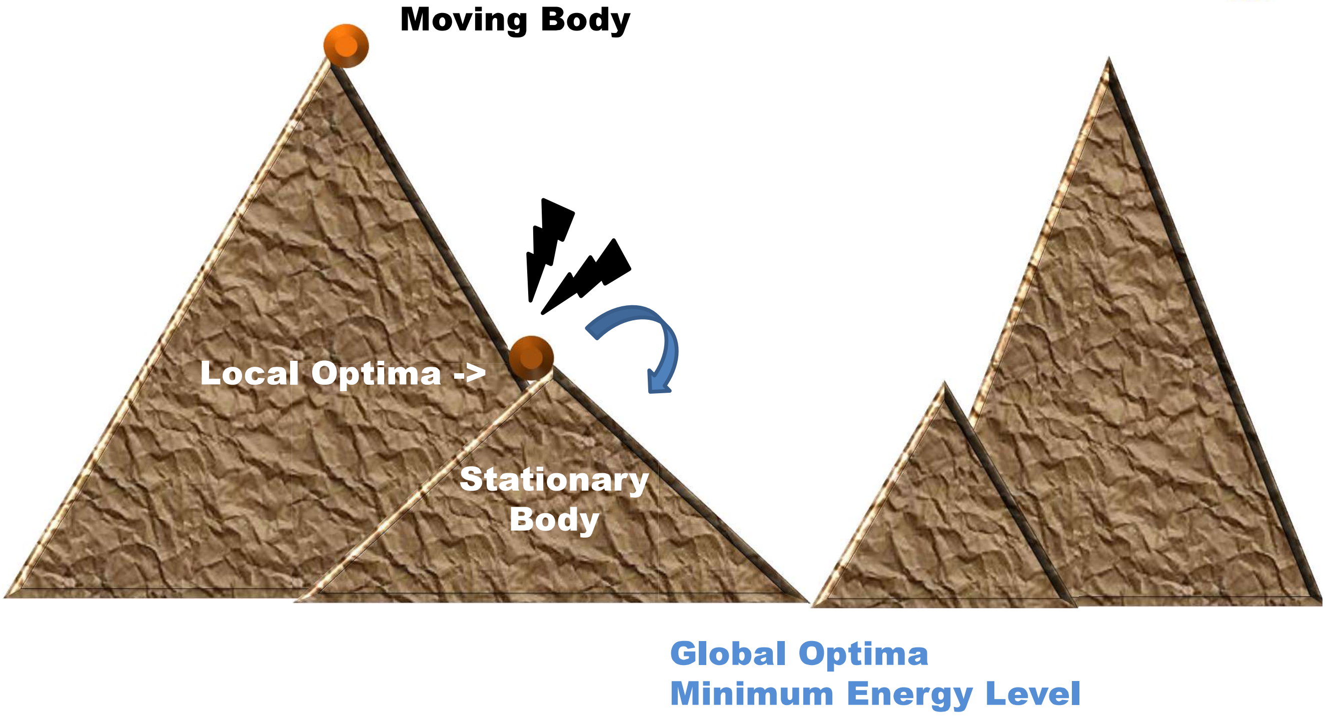
Colliding Bodies Optimization

A. Kaveh and V. Mahdavi, “Colliding bodies optimization: a novel meta-heuristic method,” Computers & Structures, Elsevier, vol. 139, pp. 18–27, 2014.

- Inspired by physical phenomenon of collision between bodies which move them to a lower energy level in the search space.
- Parameter free algorithm (no tuning of parameter is required).
- Consist of simple equations which govern the velocity and position update thus computational complexity involve is very low.

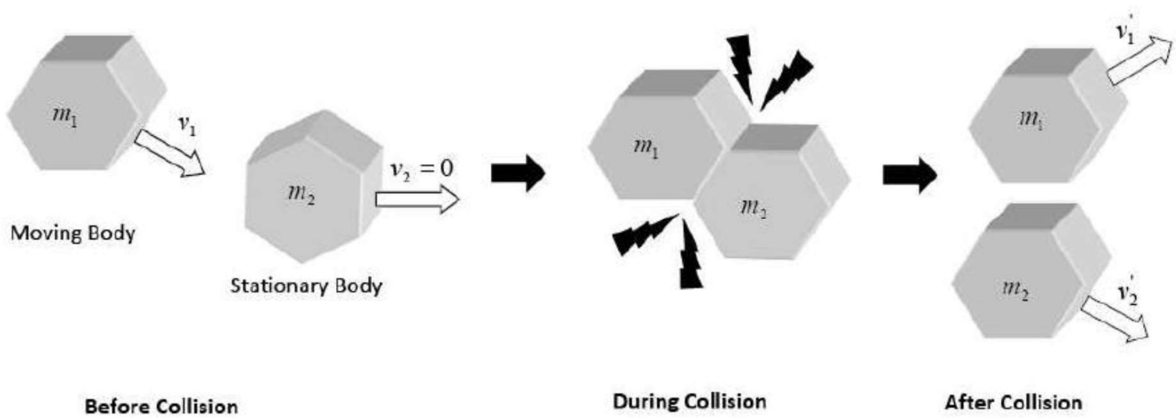
BOOK : A. Kaveh, V. R. Mahdavi, *Colliding Bodies Optimization: Extensions and Applications*, Springer Verlag, Switzerland, 2015.

Colliding Bodies Optimization



Collision based on Coefficient of Restitution ϵ

Colliding Bodies Optimization (CBO)



Conservation of Momentum $$m_1 v_1 + m_2 v_2 = m_1 v_1' + m_2 v_2'$$

Conservation of Kinetic Energy $$\frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 = \frac{1}{2} m_1 v_1'^2 + \frac{1}{2} m_2 v_2'^2 + Q$$

New velocities after collision

$$v_1' = \frac{(m_1 - \varepsilon m_2) v_1 + (m_2 + \varepsilon m_2) v_2}{m_1 + m_2} \qquad v_2' = \frac{(m_2 - \varepsilon m_1) v_2 + (m_1 + \varepsilon m_1) v_1}{m_1 + m_2}$$

Where Coefficient of Restitution (CR) is ε

Colliding Bodies Optimization (CBO)

- **Step 1 : Initialize the positions of colliding bodies :**

In CBO each solution x_i is consider as colliding body. The random initialization of initial position of K number of colliding bodies are determined with

$$x_i = x_{min} + rand\left(x_{max} - x_{min}\right), i = 1,2,\ldots,K$$

- **Step 2 : Evaluate the fitness of Colliding Bodies**

Calculate the fitness of each bodies f(x_i) and short them by best fitness value.

- **Step 3 : Calculate the mass of Colliding Bodies**

For minimization problem :

$$m_i = \frac{\frac{1}{fit(i)}}{\sum_{i=1}^{K} \frac{1}{fit(i)}}, \; i = 1,2,\ldots,K$$

For maximization problem $\frac{1}{fit(i)}$ is replaced by $fit(i)$.

BOOK : A. Kaveh, V. R. Mahdavi, *Colliding Bodies Optimization: Extensions and Applications*, Springer Verlag, Switzerland, 2015.

Colliding Bodies Optimization (CBO)

Step 4 : Initial velocities of Colliding Bodies :

Initialize half of the most fittest bodies as stationary bodies and others are moving bodies

Stationary CBs :

$$v_i = 0; \forall i = 1,2,\ldots,\frac{K}{2}$$

Moving CBs :

$$v_i = x_i - x_{i-\frac{K}{2}};$$

$$\forall i = \frac{K}{2}+1, \frac{K}{2}+2, \ldots, K$$

Step 5: Velocities of Bodies after collision :

Stationary CBs :

$$v_i' = \frac{\left(m_{i+\frac{K}{2}} + \varepsilon m_{i+\frac{K}{2}}\right) v_{i+\frac{K}{2}}}{m_i + m_{i+\frac{K}{2}}}; \forall i = 1,2,\ldots,\frac{K}{2}$$

Moving CBs :

$$v_i' = \frac{\left(m_i - \varepsilon m_{i-\frac{K}{2}}\right) v_i}{m_i + m_{i-\frac{K}{2}}}; \forall i = \frac{K}{2}+1, \frac{K}{2}+2, \ldots, K$$

Colliding Bodies Optimization (CBO)

- **Step 6 : Update the Positions of Colliding Bodies :**

Stationary CBs : $$x_i^{new} = x_i + rand * v_i' \qquad \forall i = 1,2,\ldots,\frac{K}{2}$$

Moving CBs : $$x_i^{new} = x_{i-K/2} + rand * v_i' \qquad \forall i = \frac{K}{2}+1, \frac{K}{2}+2,\ldots,K$$

Where rand is a random number uniformly distributed in the range (-1,1).

Step 7 : Calculation of Coefficient of Restitution :

We consider COR decreases linearly to zero and defined as

$$\varepsilon = 1 - \frac{iter}{iter_{\max}}$$

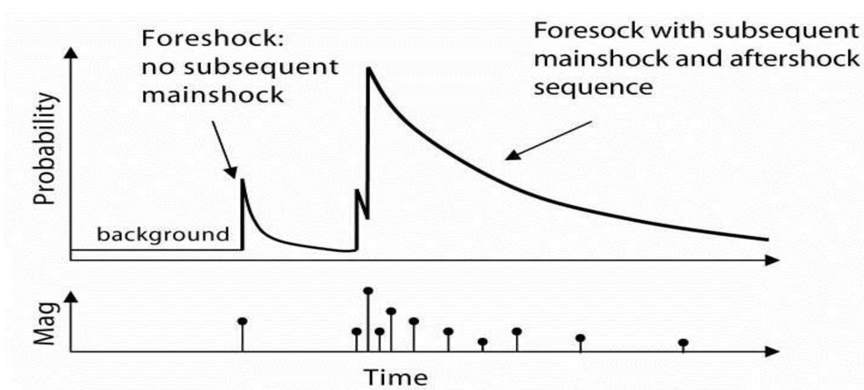
Step 8 : Termination Criteria

The **Steps 2-8** are repeated until a termination criterion such as **maximum number of iteration or get a sufficiently good fitness value** is achieved.

CBO Application to Cluster Analysis

S. J. Nanda and G. Panda. "A clustering model based on colliding bodies optimization for analysis of seismic catalog." In *2015 IEEE International Conference on Microwave, Optical and Communication Engineering (ICMOCE)*, IIT Bhubaneswar, pp. 68-71, 2015.

Problem



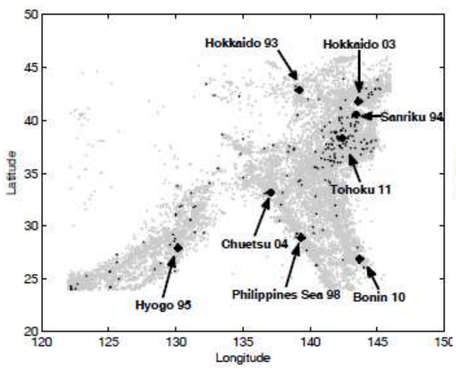
Objective

To classify the foreshocks, aftershocks and main shocks considering the time, co-ordinates and magnitude information of a synthetic / real catalogue.

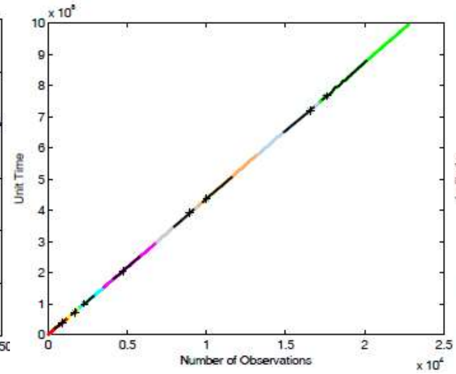
- Synthetic Catalogue : Development of the clustering model
- Real Catalogue : Validation of the model

CBO Application to Cluster Analysis

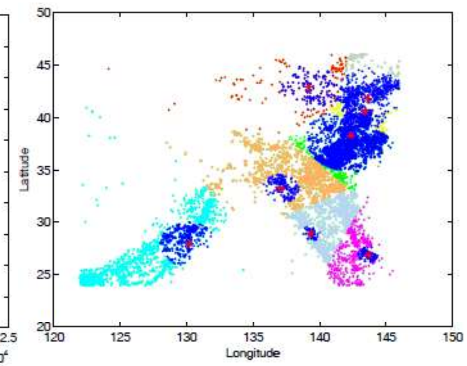
Japan Seismic Catalogue



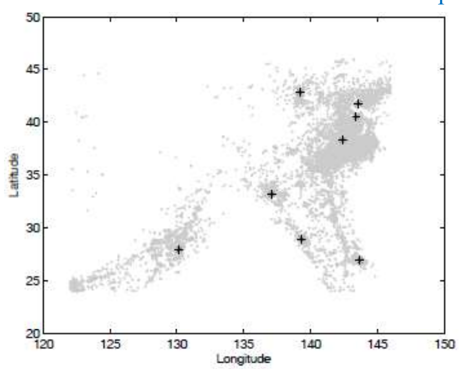
Locations of Mainshocks M > 6.8



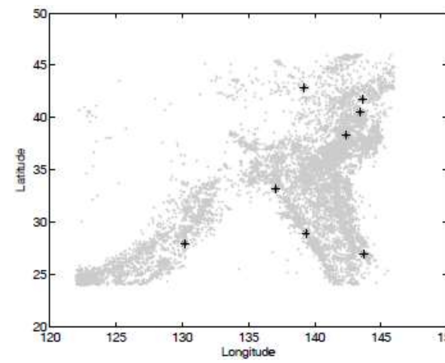
Temporal Clustering



Coordinate Clustering I

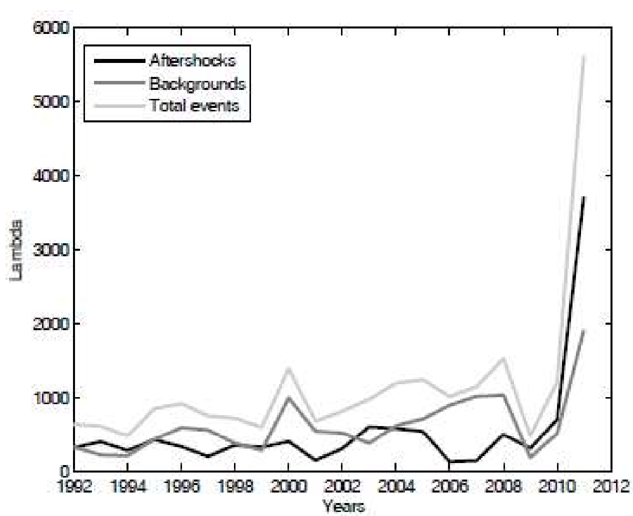


Events Classified as Aftershocks

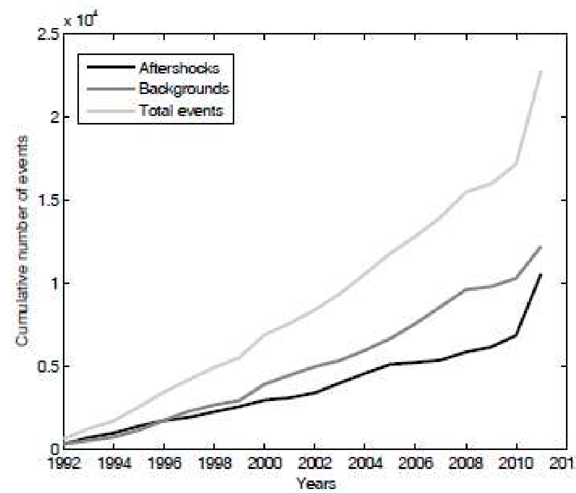


Events Classified as Backgrounds

CBO Application to Cluster Analysis



Lambda Plot for Japan



Cumulative Plot for Japan

CBO Application to Cluster Analysis

Methods	Catalogs	California	Japan	Indonesian
	Total events	111941	22794	32168
Gardner	BG events	22556	6810	7539
and	AF Clusters	7780	2040	2912
Knopoff	AF events	89385	15984	24629
Uhrhammer	BG events	51034	10807	16085
	AF Clusters	7688	1004	1757
	AF events	60907	11987	16083
Proposed	BG events	68217	12212	19861
model	AF events	43724	10582	12307

- It is observed that the proposed method does not over classify events as aftershocks, as occurs in both the standard methods.

THANK YOU