

Genetic Algorithm based Multi-objective Optimization

Prof. Ganapati Panda, FNAE, FNASc.

IIT Bhubaneswar

“Multiobjective optimization is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints.”

Examples of multi-objective optimization problems:-

- ❖ Maximizing profit and minimizing the cost of a product.
- ❖ Maximizing performance and minimizing fuel consumption of a vehicle.
- ❖ Minimizing weight while maximizing the strength of a particular component.

Difference

Single Objective Optimization

- ❖ Optimize only one objective function
- ❖ Single optimal solution
- ❖ Maximum/Minimum fitness value is selected as the best solution.

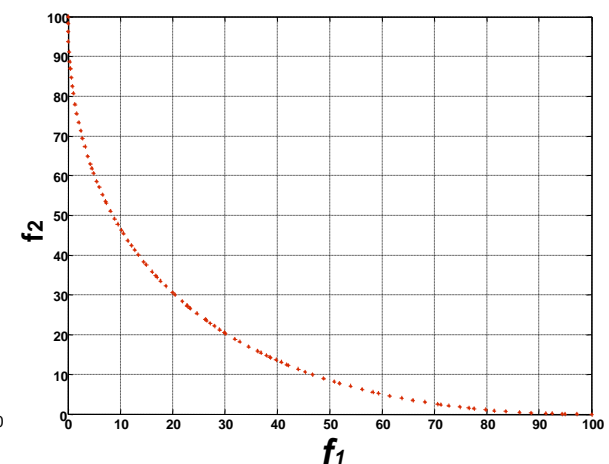
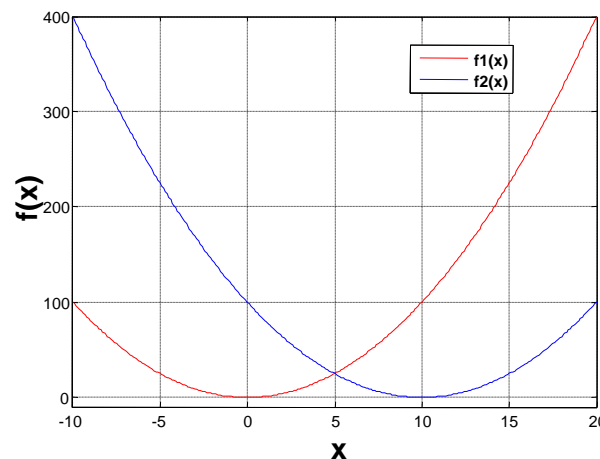
Minimize

$$f_1(x) = x^2, \quad f_2(x) = (x - 10)^2;$$

where $-10 < x < 20$

Optimal solution:-

$$x \in [0, 10]$$



Multiobjective Optimization

- ❖ Optimize two or more than two objective functions
- ❖ Set of optimal solutions
- ❖ Comparison of solutions by
 - Domination
 - Non-domination

Standard Approach : Weighted Sum of Objective Functions

$$\text{Minimize} - \{f_1(\vec{x}), f_2(\vec{x}), \dots, f_m(\vec{x})\}$$

Formulate as a single objective with weighted sum of all objective functions -

$$g(\vec{x}) = \lambda_1 f_1(\vec{x}) + \lambda_2 f_2(\vec{x}) + \dots + \lambda_m f_m(\vec{x})$$

where $\lambda_1, \lambda_2, \dots, \lambda_m$ are weights values & $\lambda_1 + \lambda_2 + \dots + \lambda_m = 1$ and m represents the number of objective functions.

Limitations:

- ✿ Result depends on weights.
- ✿ Some solutions may be missed.
- ✿ Multiple runs of the algorithm are required in order to get the whole range of solutions.
- ✿ Difficult to select proper combination of weights.
- ✿ Combining objectives loses information and predetermines trade-offs between objectives.

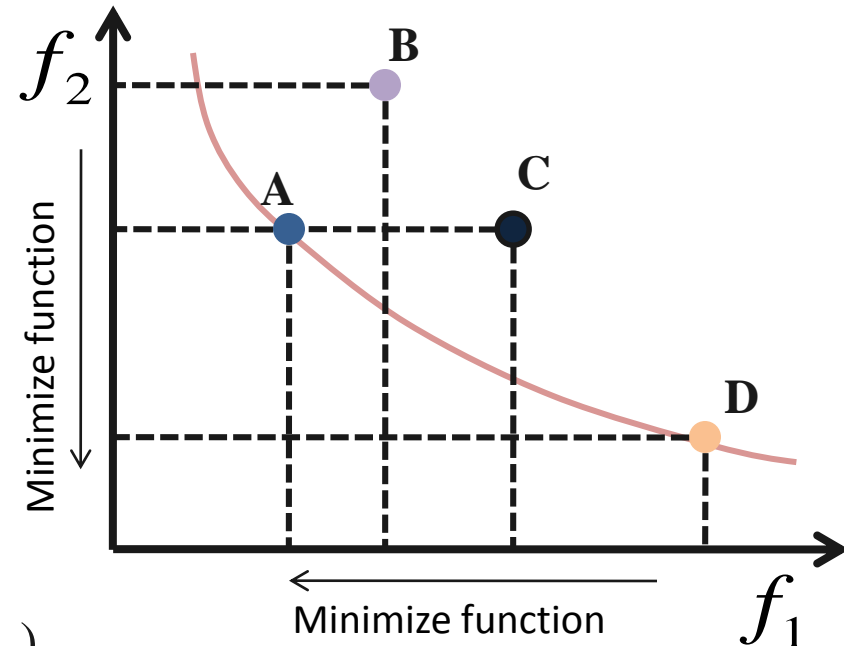
Definitions

Domination :

One solution is said to dominate another if it is better in all objectives.

Non-Domination [Pareto points] :

A solution is said to be non-dominated if it is better than other solutions in at least one objective.



- ✓ A dominates B (better in both f_1 and f_2)
- ✓ A dominates C (same in f_1 but better in f_2)
- ✓ A does not dominate D (non-dominated points)
- ✓ A and D are in the “Pareto optimal front”
- ✓ These non-dominated solutions are called Pareto optimal solutions.
- ✓ This non-dominated curve is said to be **Pareto front**.

Pareto Optimal

A vector variable $\vec{x}^* \in \Omega$ is Pareto optimal if for every $\vec{x} \in \Omega$ and $I = \{1, 2, \dots, k\}$ either

$$\forall_{i \in I} \quad f_i(\vec{x}^*) \leq f_i(\vec{x})$$

or, there is at least one $i \in I$ such that

$$\exists_{i \in I} \quad f_i(\vec{x}^*) < f_i(\vec{x})$$

where $\vec{x} = [x_1, x_2, \dots, x_n]^T$ is the vector of decision variables, $\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \dots, f_k(\vec{x})]^T$ is the vector of objective functions, Ω is the feasible region ($\Omega \in S$), where S represents the whole search space.

Pareto Optimal Set

For a given MOP $\vec{f}(\vec{x})$ the Pareto optimal set is (p^*) defined as

$$p^* := \{ \vec{x} \in \Omega \mid \neg \exists \vec{x}' \in \Omega \ \vec{f}(\vec{x}') \leq \vec{f}(\vec{x}) \}$$

Pareto Front

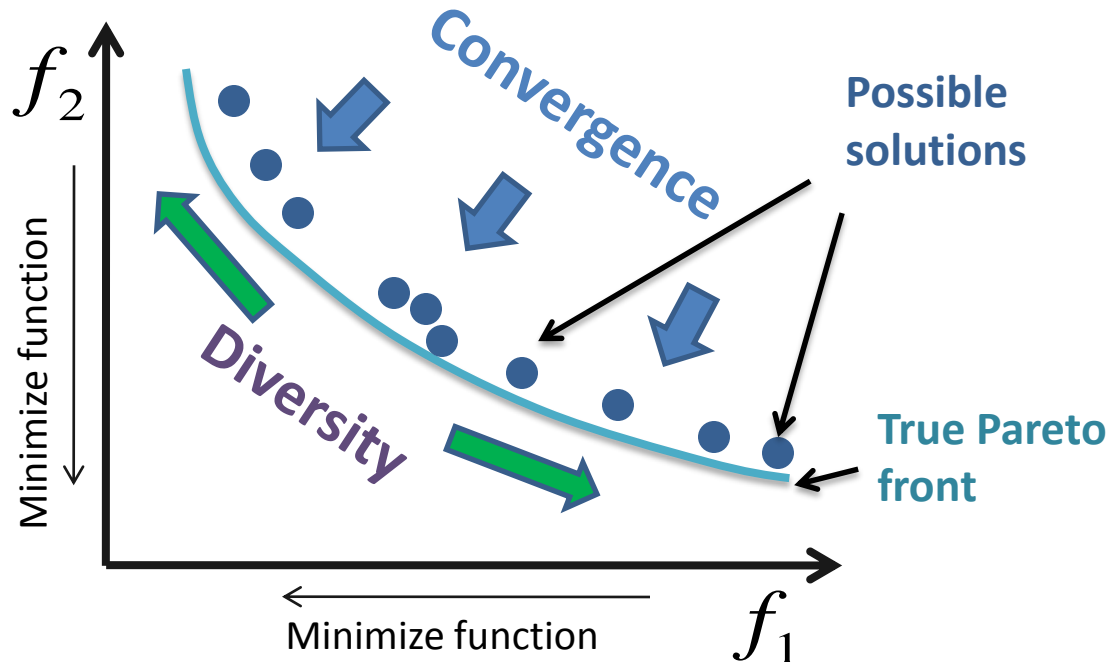
For a given MOP $\vec{f}(\vec{x})$ and Pareto optimal set p^* , the Pareto front pf^* is defined as

$$pf^* := \{ \vec{f} = (f_1(\vec{x}), f_2(\vec{x}), \dots, f_k(\vec{x})) \mid \vec{x} \in p^* \}$$

Desirable MOEA features

Best Solutions

- Lie on true Pareto front
- They are uniformly distributed on the front



Aim:

- To achieve convergence to Pareto optimal front
- To achieve diversity (representation of the entire Pareto optimal front)

Non Dominated Sorting based Genetic Algorithm II (NSGA- II)

- ❑ Developed by Prof. K. Deb at **Kanpur Genetic Algorithms Laboratory** (2002)
- ❑ Famous for *Fast non-dominated search*
- ❑ Fitness assignment - *Ranking based* on non-domination sorting
- ❑ Diversity mechanism is based on *Crowding distance*
- ❑ Uses *Elitism*

Initialize Population

Minimize

$$f_1(x) = x^2, f_2(x) = (x - 2)^2$$

where $-5 \leq x \leq 5$

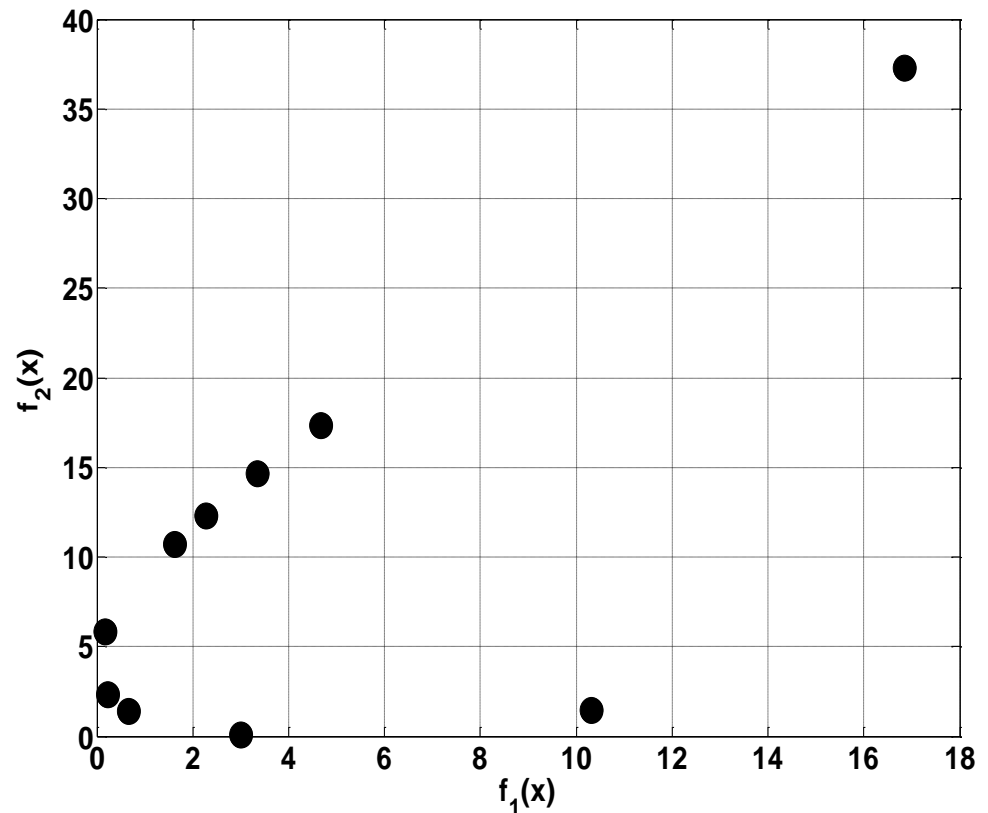
- Search space is of single dimension (given).
- Objective space is of two dimension (given).
- Let population size = 10
- Initialize population with 10 chromosomes having single dimensioned real value.
- These values are randomly distributed in between [-5,5].

x
0.4678
1.7355
0.8183
-0.414
3.2105
-1.272
-1.508
-1.832
-2.161
-4.105

Evaluate Fitness values

- Find out all objective functions values (fitness values) for all chromosomes.

x	$f_1(x)$	$f_2(x)$
-0.414	0.171	5.829
0.467	0.218	2.347
0.818	0.669	1.396
1.735	3.011	0.07
3.210	10.308	1.465
-1.272	1.618	10.708
-1.508	2.275	12.308
-1.832	3.355	14.682
-2.161	4.671	17.317
-4.105	16.854	37.275



Fast Non-domination Sorting

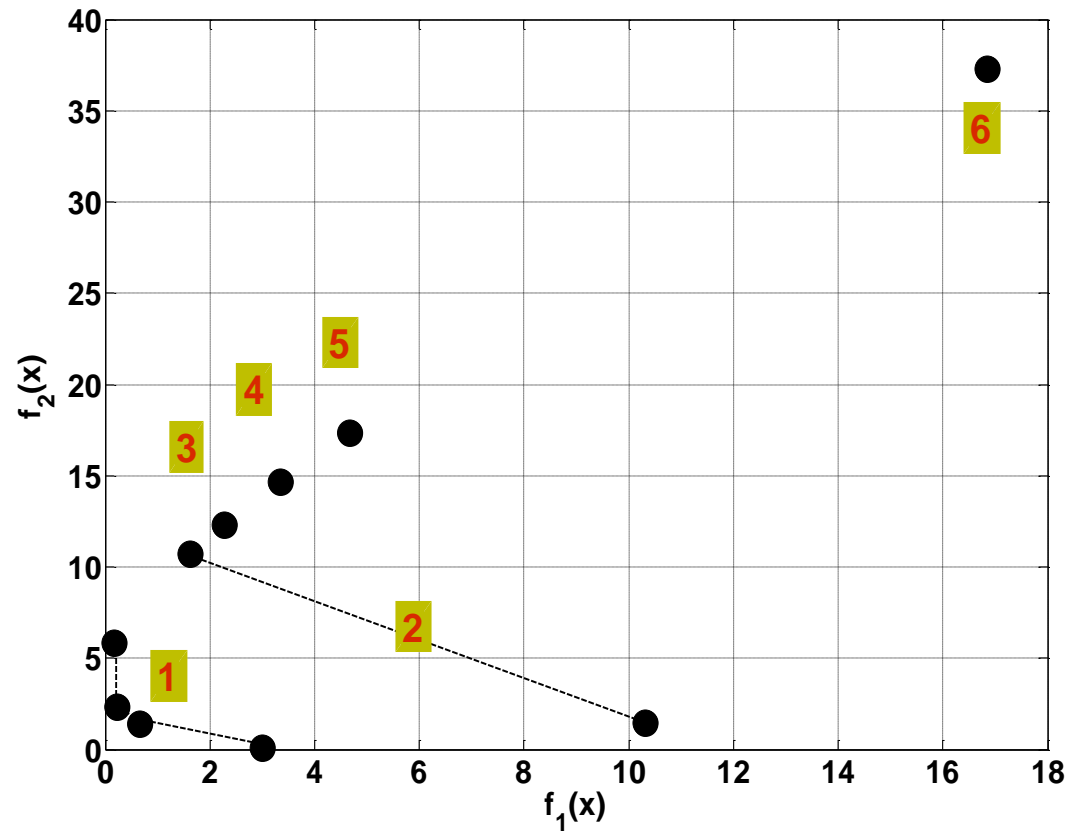
- Assigning the rank to each individual of the population.
- Rank based on the *non-domination sorting* (front wise).
- It helps in selection and sorting.

x	$f_1(x)$	$f_2(x)$	Reference chromosome	Dominated chromosomes	Counter	Rank
-0.414	0.171	5.829				1
0.467	0.218	2.347				1
0.818	0.669	1.396				1
1.735	3.011	0.07				1
3.210	10.308	1.465				2
-1.272	1.618	10.708				2
-1.508	2.275	12.308				3
-1.832	3.355	14.682				4
-2.161	4.671	17.317				5
-4.105	16.854	37.275				6

$x^1 = \{\dots\} \Rightarrow n^1 = 0$
$x^2 = \{\dots\} \Rightarrow n^2 = 0$
$x^3 = \{\dots\} \Rightarrow n^3 = 0$
$x^4 = \{\dots\} \Rightarrow n^4 = 0$
$x^5 = \{\underline{x^3}, \underline{x^4}\} \Rightarrow n^5 = 1$
$x^6 = \{\underline{x^1}, \underline{x^2}, \underline{x^3}\} \Rightarrow n^6 = 1$
$x^7 = \{\underline{x^1}, \underline{x^2}, \underline{x^3}, \underline{x^6}\} \Rightarrow n^7 = 2$

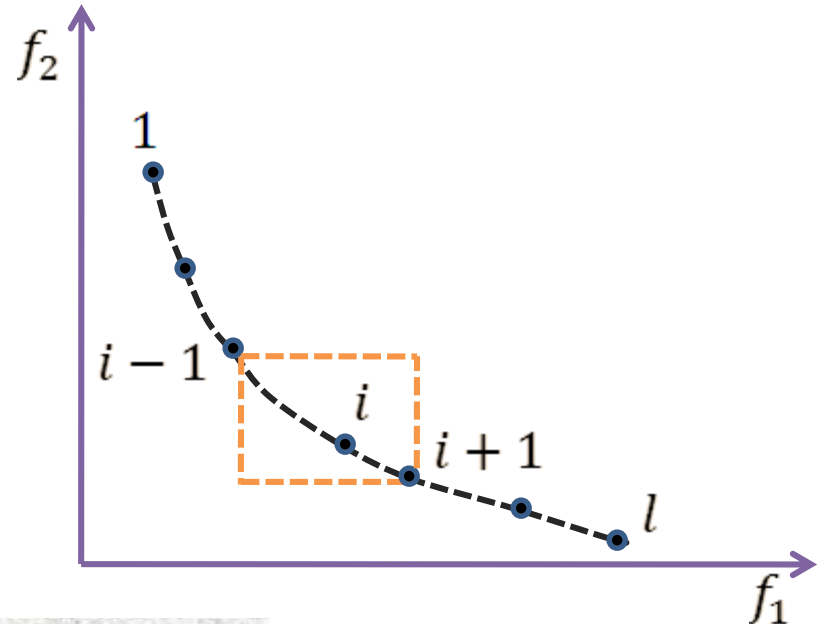
Fast Non-domination Sorting

x	$f_1(x)$	$f_2(x)$	Rank
-0.414	0.171	5.829	1
0.467	0.218	2.347	1
0.818	0.669	1.396	1
1.735	3.011	0.07	1
3.210	10.308	1.465	2
-1.272	1.618	10.708	2
-1.508	2.275	12.308	3
-1.832	3.355	14.682	4
-2.161	4.671	17.317	5
-4.105	16.854	37.275	6



Crowding Distance Assignment

- To get an estimate of density of solutions surrounding a particular solution in population.
- Choose individuals having large crowding distance.
- Help for obtaining uniformly distribution.



$$l_{C.D.} = l_{C.D.} = \infty,$$

$$i_{C.D.} = \sum_m \left(\frac{f[i+1]_m - f[i-1]_m}{f_m^{\max} - f_m^{\min}} \right)$$

where
 $i = 2, 3, \dots, (l-1)$

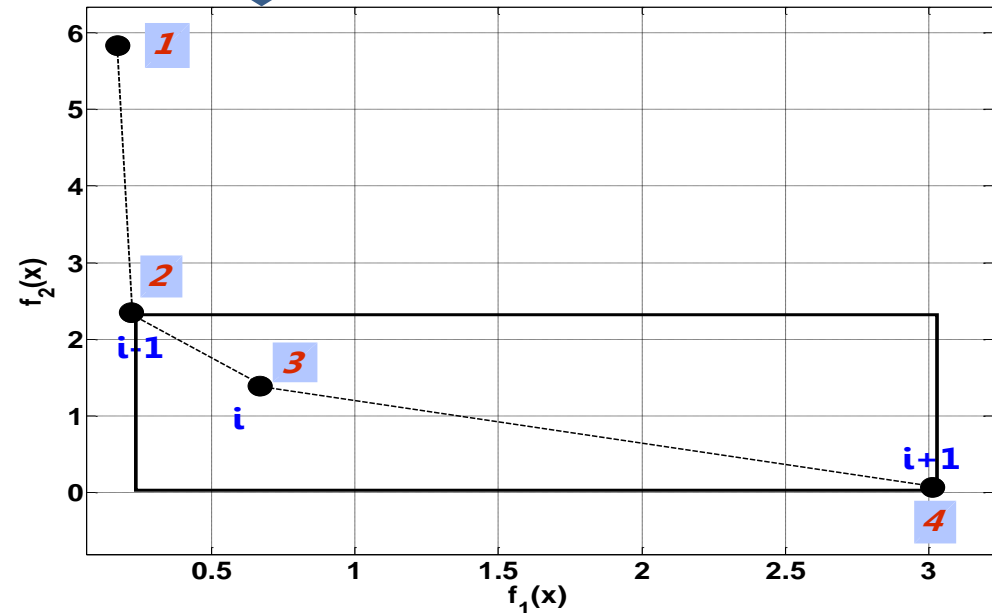
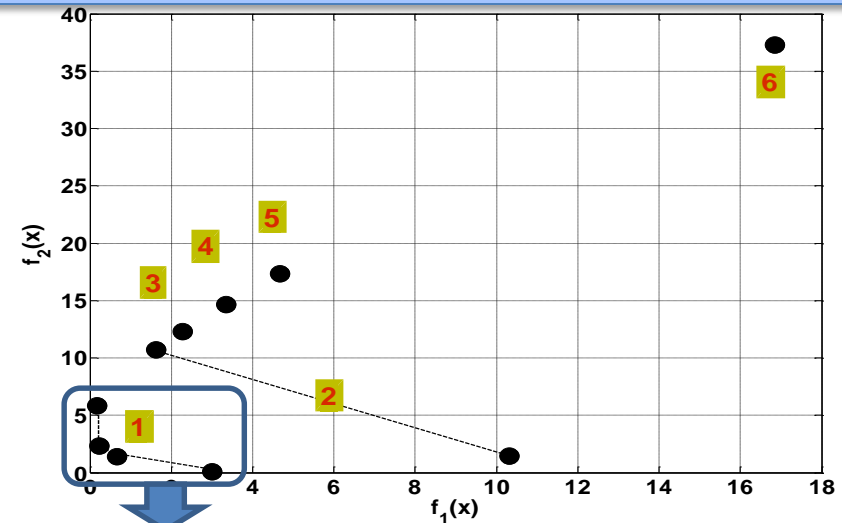
where $f[i]_m$ represent m^{th} objective function value of i^{th} solution.

and f_m^{\max} is the maximum value of function \vec{f}_m in the Pareto front.

Crowding Distance Assignment

- Crowding distance can be calculated for all chromosomes of same Pareto front.

x	$f_1(x)$	$f_2(x)$	Rank	C.D.
-0.414	0.171	5.829	1	∞
0.467	0.218	2.347	1	0.945
0.818	0.669	1.396	1	1.378
1.735	3.011	0.07	1	∞
3.210	10.308	1.465	2	∞
-1.272	1.618	10.708	2	∞
-1.508	2.275	12.308	3	∞
-1.832	3.355	14.682	4	∞
-2.161	4.671	17.317	5	∞
-4.105	16.854	37.275	6	∞



Selection is the stage of a genetic algorithm in which individual are chosen from a population for later breeding (recombination or crossover).

Crowding operator based sorting

The crowding operator \leq_n guides the selection process at the various stages of the algorithm toward a uniformly spread-out Pareto optimal front.

$$i \leq_n j \quad \left(i_{rank} < j_{rank} \right)$$

or

$$\left(i_{rank} = j_{rank} \right) \text{ and } \left(i_{C.D.} > j_{C.D.} \right)$$

where i_{rank} shows non-domination rank & $i_{C.D.}$ is crowding distance of i^{th} individual.

Tournament Selection

Runs a 'tournament' among a few individuals chosen at random from the population and selects the winner (the one with the best fitness) for crossover.

- In tournament selection, a number *Tour size* of individuals is chosen randomly from the population and the best individual from this group is selected as parent. (Based on the crowding operator)

x	$f_1(x)$	$f_2(x)$	Rank	C.D.
0.818	0.669	1.396	1	1.378
-1.508	2.275	12.30	3	∞

$$1_{rank} < 2_{rank}$$



0.818	0.669	1.396	1	1.378
-------	-------	-------	---	-------

x	$f_1(x)$	$f_2(x)$	Rank	C.D.
0.467	0.218	2.347	1	0.945
0.818	0.669	1.396	1	1.378

$$1_{rank} = 2_{rank} \quad 1_{C.D.} < 2_{C.D.}$$



0.818	0.669	1.396	1	1.378
-------	-------	-------	---	-------

Crossover

***Crossover** is a genetic operator that combines (mates) two individuals (parents) to produce two new individuals (Childs).*

- The idea behind crossover is that the new chromosome may be better than both of the parents if it takes the best characteristics from each of the parents.

Simulated Binary Crossover

$$b = \begin{cases} (2 * r)^{\left(\frac{1}{\mu+1}\right)} & \text{if } r \leq 0.5 \\ \left(\frac{1}{2 * (1-r)}\right)^{\left(\frac{1}{\mu+1}\right)} & \text{if } r > 0.5 \end{cases}$$

where

- r is a random number $\{0,1\}$
- μ is a crossover operator
- j represent dimension of individual.

$$child_1(j) = \frac{1}{2} \left((1+b) * parent_1(j) + (1-b) * parent_2(j) \right)$$

$$child_2(j) = \frac{1}{2} \left((1-b) * parent_1(j) + (1+b) * parent_2(j) \right)$$

Mutation is a genetic operator that alters one or more gene values in a chromosome from its initial state.

- Mutation is an important part of the genetic search as helps to prevent the population from stagnating at any local optima.

Polynomial Mutation

$$d = \begin{cases} (2 * r)^{\left(\frac{1}{\eta+1}\right)} - 1 & \text{if } r \leq 0.5 \\ 1 - (2 * (1 - r))^{\left(\frac{1}{\eta+1}\right)} & \text{if } r > 0.5 \end{cases}$$

$$\text{child}(j) = \text{parent}(j) + d$$

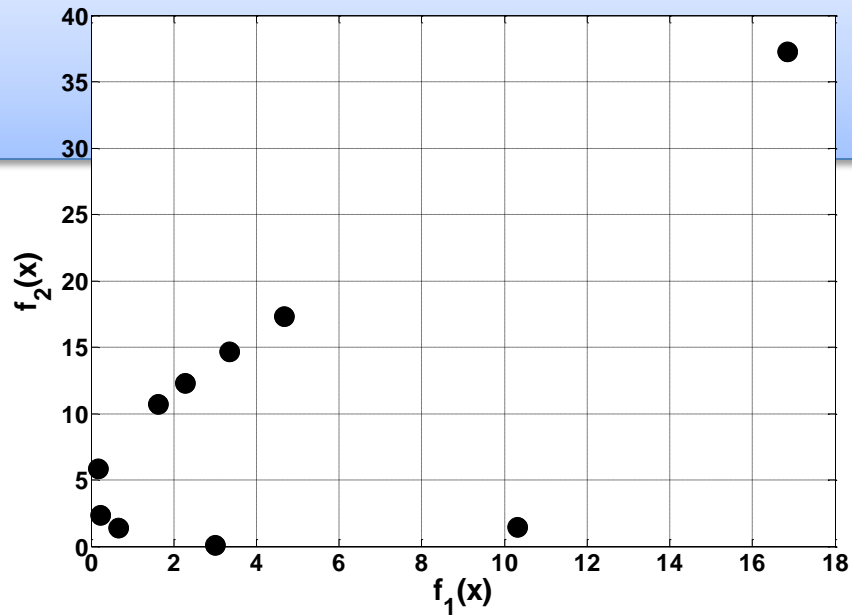
where

- r is a random number $\{0,1\}$
- η is a mutation operator
- j represent dimension of individual.

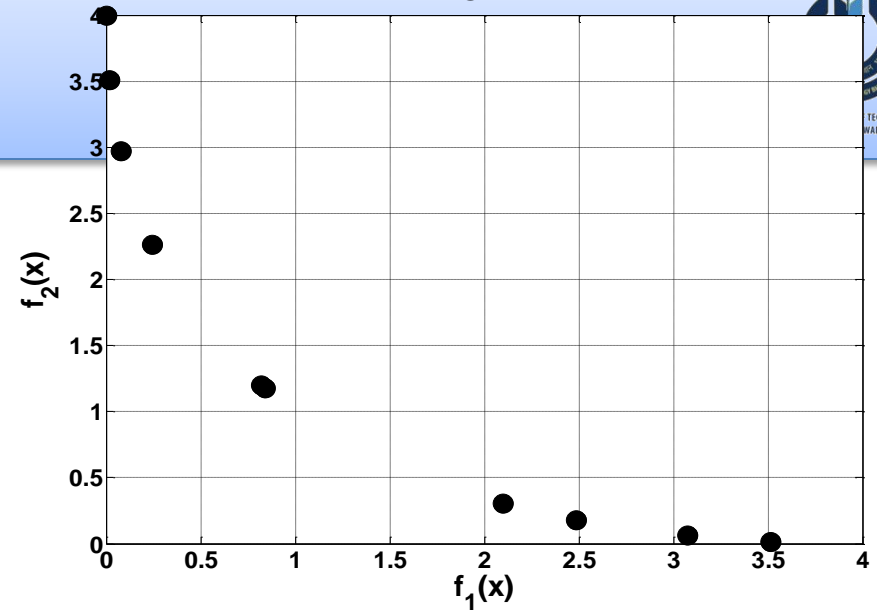
Crowding distance sorting



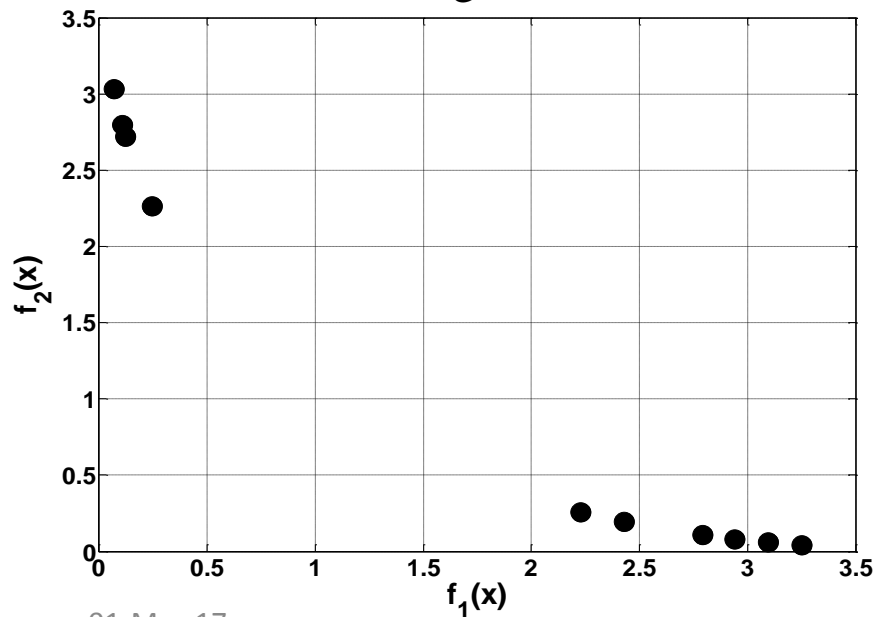
Initial State



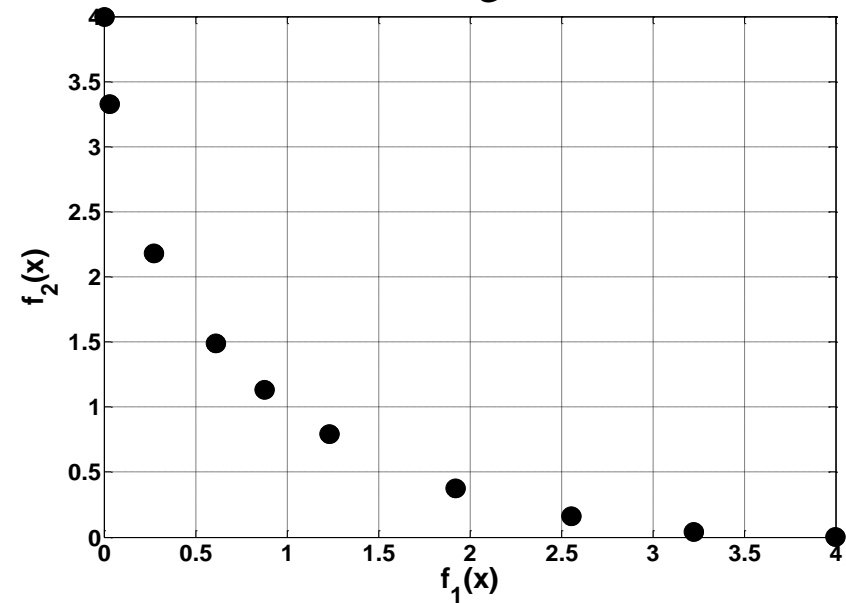
After 20 generation



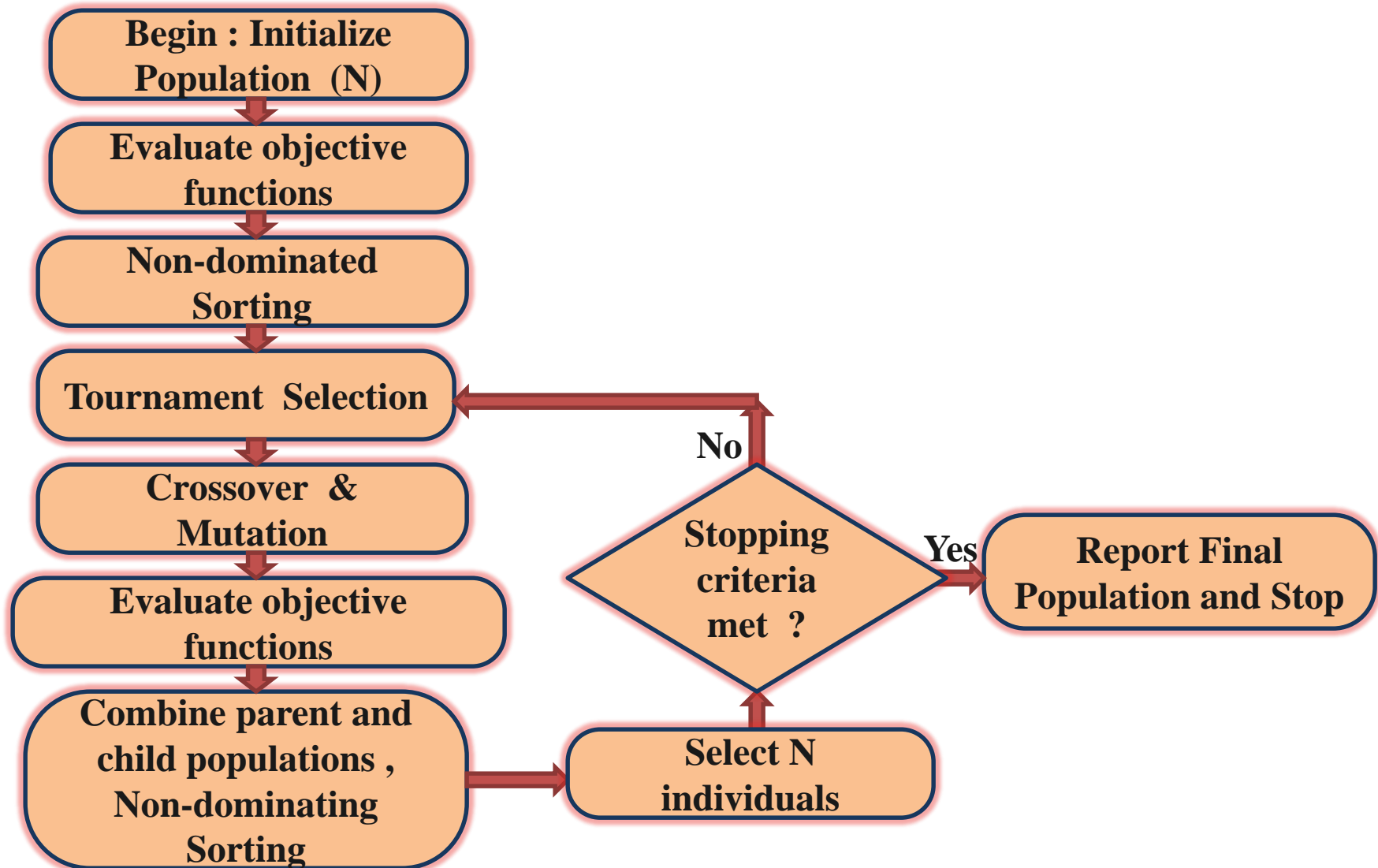
After 10 generation



After 40 generation



Flowchart of NSGA-II



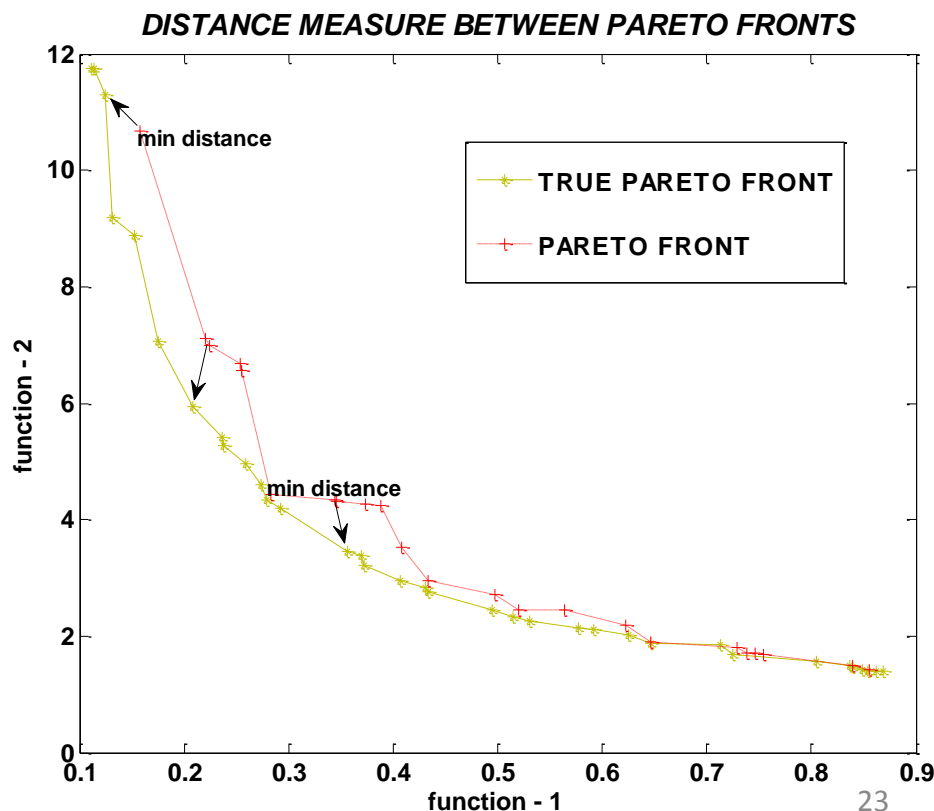
Performance Measures

There are two main goals in a multi-objective optimization:

- 1) *Convergence to the Pareto-optimal set*
- 2) *Maintenance of diversity in solutions of the Pareto-optimal set.*

Convergence metric

$$\gamma = avg \left(\sum_N (\min_distance) \right)$$

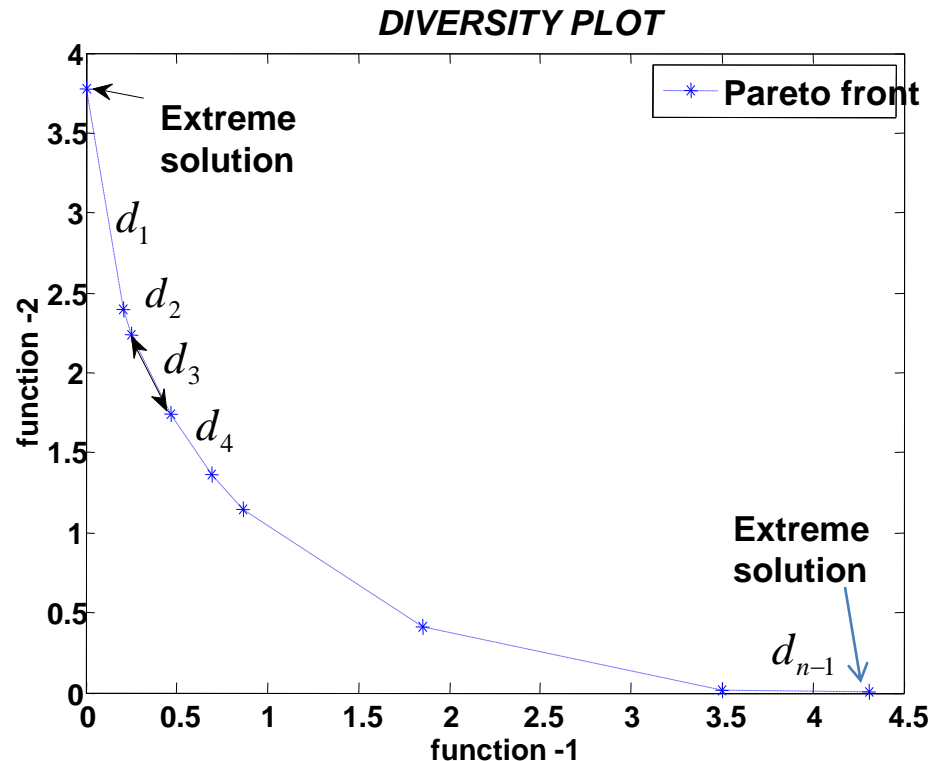


Cont....

Non uniformity in the distribution,

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |(d_i - \bar{d})|}{d_f + d_l + (N - 1)\bar{d}}$$

➤ If distance between the solutions d_i is equal to average distance \bar{d} , that gives uniformly distribution.



The parameters d_f and d_l are the Euclidean distances between the extreme solutions of true Pareto front and the boundary solutions of the obtained non-dominated set.

The parameter \bar{d} is the average of all distances, $d_i, i = 1, 2, \dots, N - 1$ assuming that there are N solutions on the best non-dominated front.

Simulation and Results

NSGA II parameters

- ✓ Population (N) = 100
- ✓ Crossover Probability (P_c) = 0.9
- ✓ Mutation Probability (P_m) = 0.1
- ✓ Distribution index for crossover (μ) = 20
- ✓ Distribution index for mutation (η) = 20
- ✓ Tour size (selection) = 2

MOPSO parameters

- ❖ Population = 100 particles
- ❖ Repository (Archive) size = 100 particles
- ❖ Mutation rate = 0.5
- ❖ Divisions for Archive Grid = 30

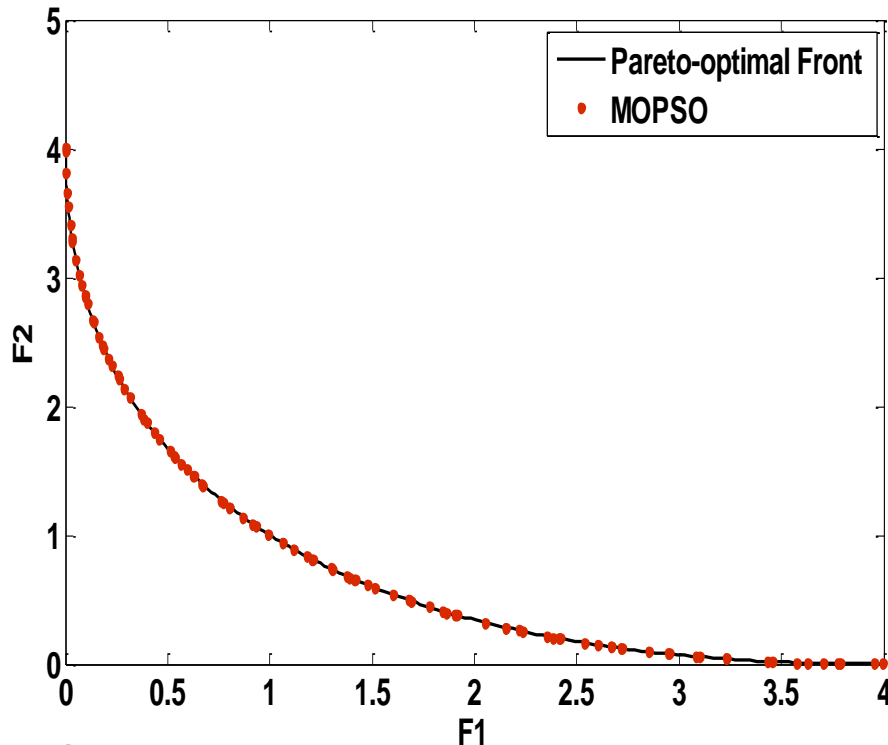
Implementation use real numbers representation.

- * These parameters were kept in all test functions optimization.
- * Only changed the total number of fitness function evaluations.

Test Problem : SCH

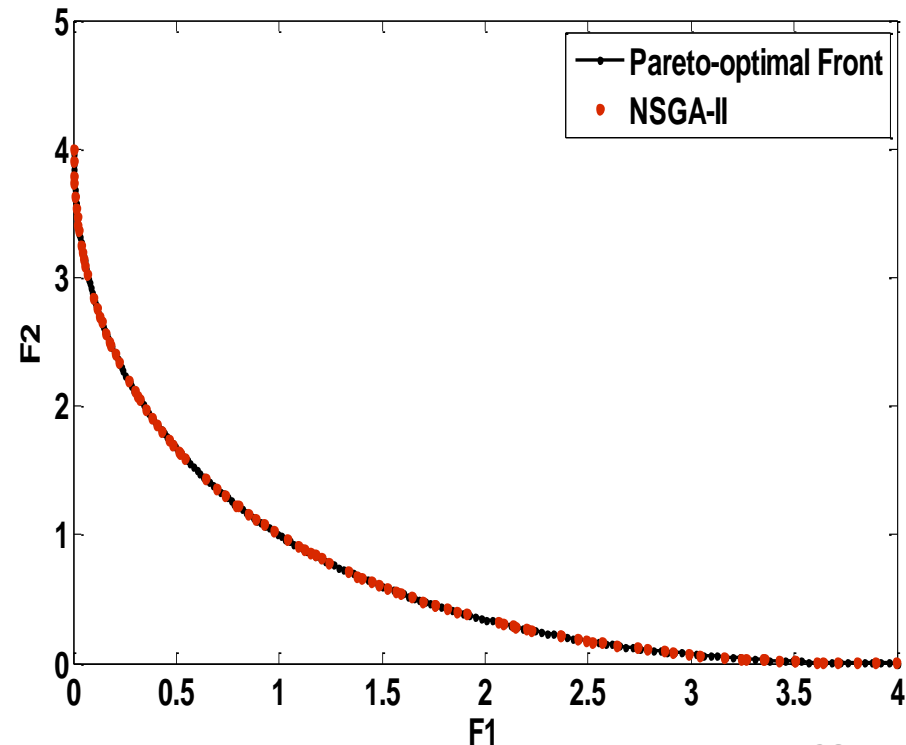
$$SCH_{Min} : \begin{cases} f_1(\mathbf{x}) = x^2 \\ f_2(\mathbf{x}) = (x - 2)^2 \\ -1000 \leq x \leq 1000 \end{cases}$$

Nondominated solutions with MOPSO on SCH



01-Mar-17

Nondominated solutions with NSGA-II on SCH



26

Comparison for SCH Test Function

Convergence Metric	NSGA-II	MOPSO
Best	0.0148	0.0093
Worst	0.9578	0.1569
Mean	0.2096	0.0259

Tab. 1: Results of the Convergence Metric for the SCH Test Function

Diversity Metric	NSGA-II	MOPSO
Best	0.5104	0.6947
Worst	0.7904	1.3575
Mean	0.6425	0.8582

Tab. 2: Results of the Diversity Metric for the SCH Test Function

* Total number of fitness evaluations was set to 10,000.

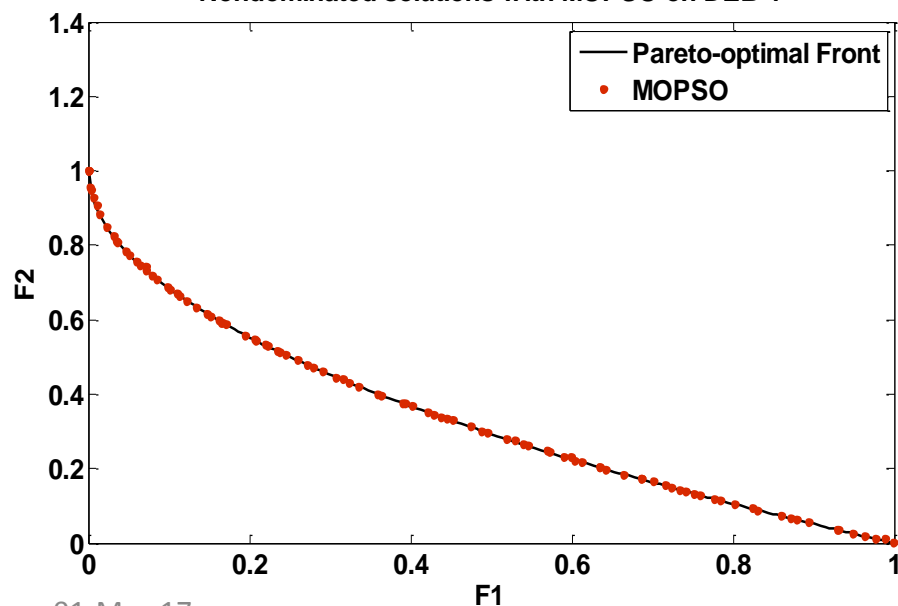
Test Problem : DEB-1

$$DEB - 1_{Min} : \begin{cases} f_1(\mathbf{x}) = x_1 \\ f_2(\mathbf{x}) = \frac{g(x_2)}{x_1} \end{cases}$$

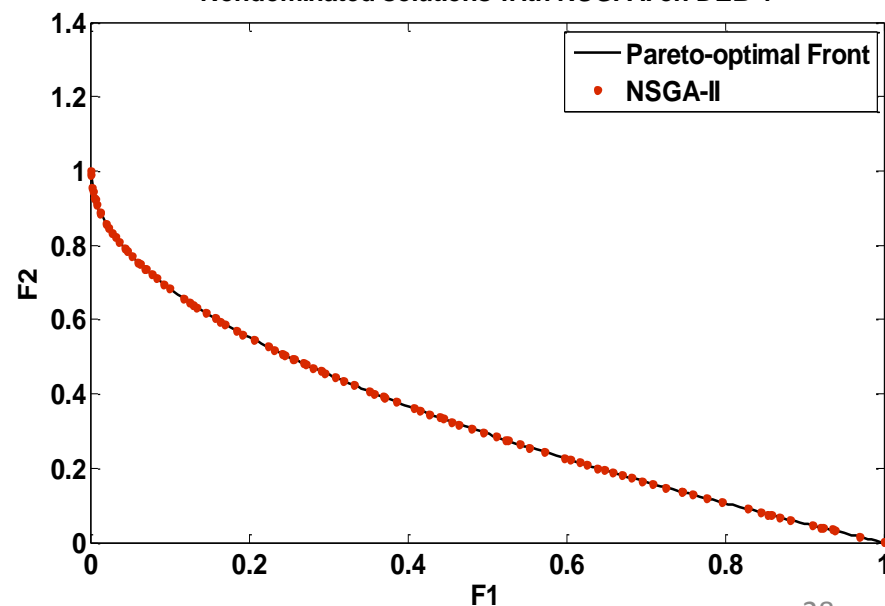
$$\text{where } g(x_2) = 2.0 - \exp\left\{-\left(\frac{(x_2 - 0.2)}{0.004}\right)^2\right\} - 0.8 \exp\left\{-\left(\frac{(x_2 - 0.6)}{0.4}\right)^2\right\}$$

$$0.1 \leq x_1, x_2 \leq 1.0$$

Nondominated solutions with MOPSO on DEB-1



Nondominated solutions with NSGA-II on DEB-1



Comparison for DEB-1 Test Function

Convergence Metric	NSGA-II	MOPSO
Best	0.0066	0.0070
Worst	0.5140	0.1664
Mean	0.0078	0.0079

Tab. 1: Results of the Convergence Metric for the DEB-1 Test Function

Diversity Metric	NSGA-II	MOPSO
Best	0.3467	0.5112
Worst	0.5140	0.7168
Mean	0.4243	0.5938

Tab. 2: Results of the Diversity Metric for the DEB-1 Test Function

* Total number of fitness evaluations was set to 15,000.

Test Problem : DEB-2

$$DEB - 2_{Min} : \begin{cases} f_1(\mathbf{x}) = x_1 \\ f_2(\mathbf{x}) = g(\mathbf{x}) \cdot h(\mathbf{x}) \end{cases}$$

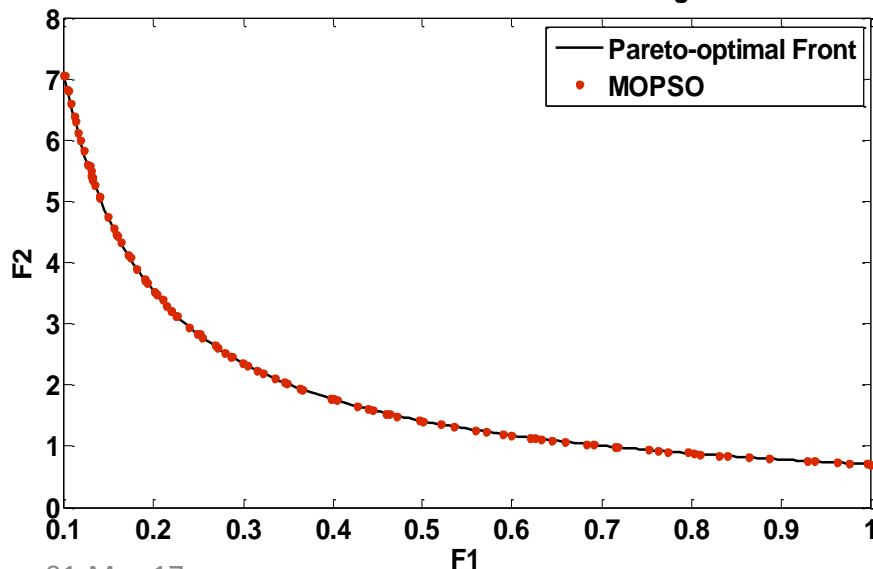
where

$$g(\mathbf{x}) = 11 + x_2^2 - 10\cos(2\pi x_2)$$

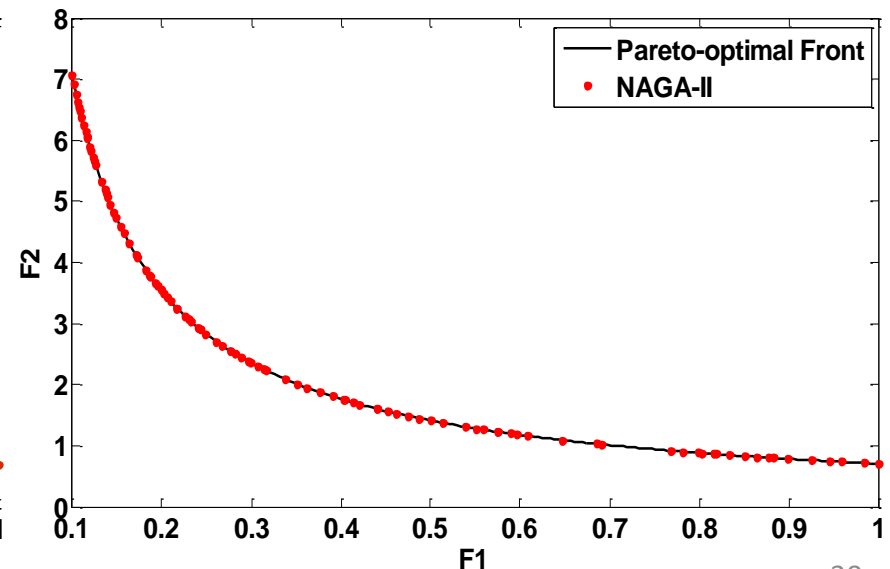
$$h(\mathbf{x}) = \begin{cases} 1 - \sqrt{\frac{f_1(\mathbf{x})}{g(\mathbf{x})}} & \text{if } f_1(\mathbf{x}) \leq g(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases}$$

$$0 \leq x_1 \leq 1, \quad -30 \leq x_2 \leq 30$$

Nondominated solutions with MOPSO using DEB-2



Nondominated solutions with NSGA-II on DEB-2



Comparison for DEB-2 Test Function

Convergence Metric	NSGA-II	MOPSO
Best	0.0449	0.0515
Worst	0.0559	0.0725
Mean	0.0516	0.0608

Tab. 1: Results of the Convergence Metric for the DEB-2 Test Function

Diversity Metric	NSGA-II	MOPSO
Best	0.7248	0.6800
Worst	0.7939	0.7582
Mean	0.7597	0.7193

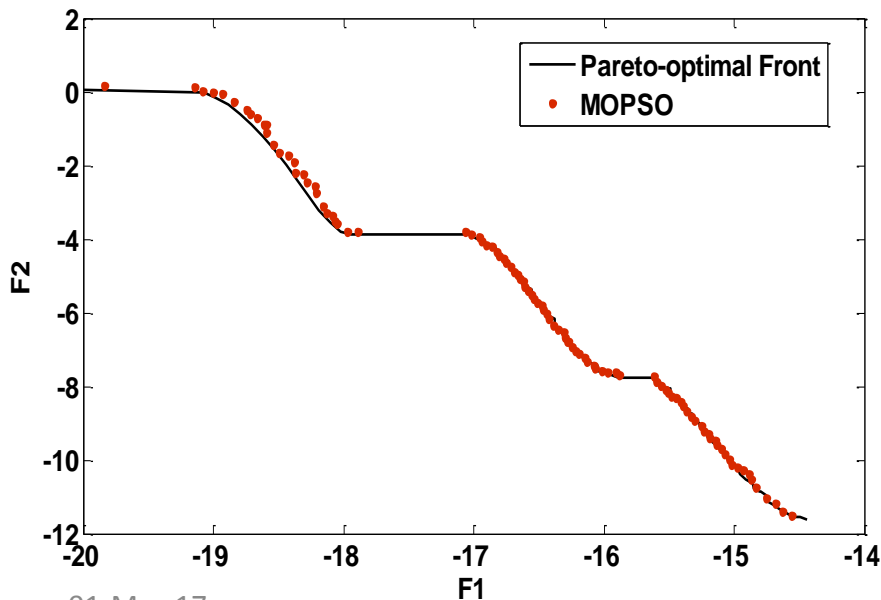
Tab. 2: Results of the Diversity Metric for the DEB-2 Test Function

* Total number of fitness evaluations was set to 25,000.

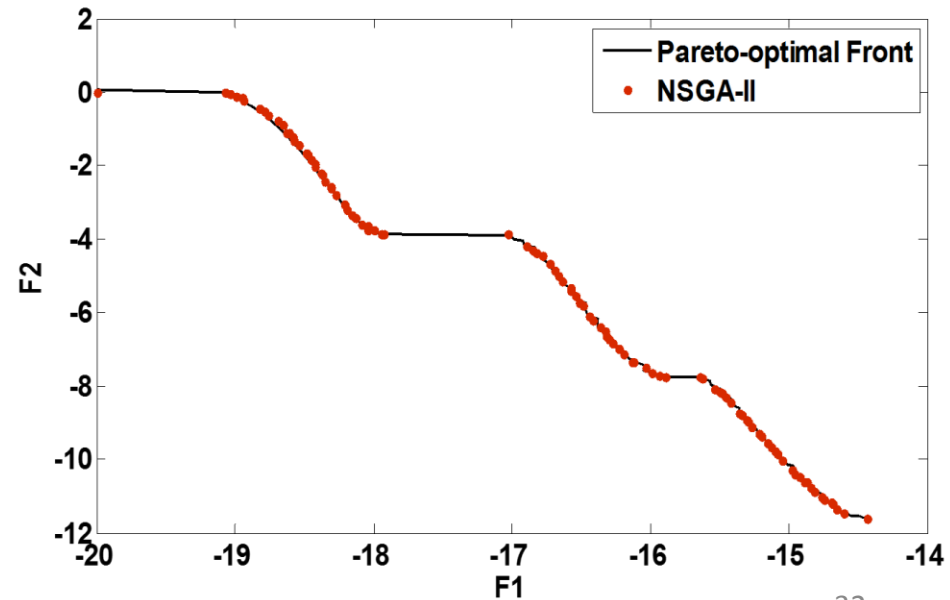
Test Problem : KUR

$$KUR_{Min} : \begin{cases} f_1(\mathbf{x}) = \sum_{i=1}^2 \left[-10 \exp \left(-0.2 \sqrt{x_i^2 + x_{i+1}^2} \right) \right] \\ f_2(\mathbf{x}) = \sum_{i=1}^3 [|x_i|^{0.8} + 5 \sin(x_i^3)] \\ -5 \leq x_i \leq 5, \quad i = 1, 2, 3 \end{cases}$$

Nondominated solutions with MOPSO on KUR



Nondominated solutions with NSGA-II on KUR



Comparison for KUR Test Function

Convergence Metric	NSGA-II	MOPSO
Best	0.0021	0.0021
Worst	0.0041	0.0034
Mean	0.0028	0.0026

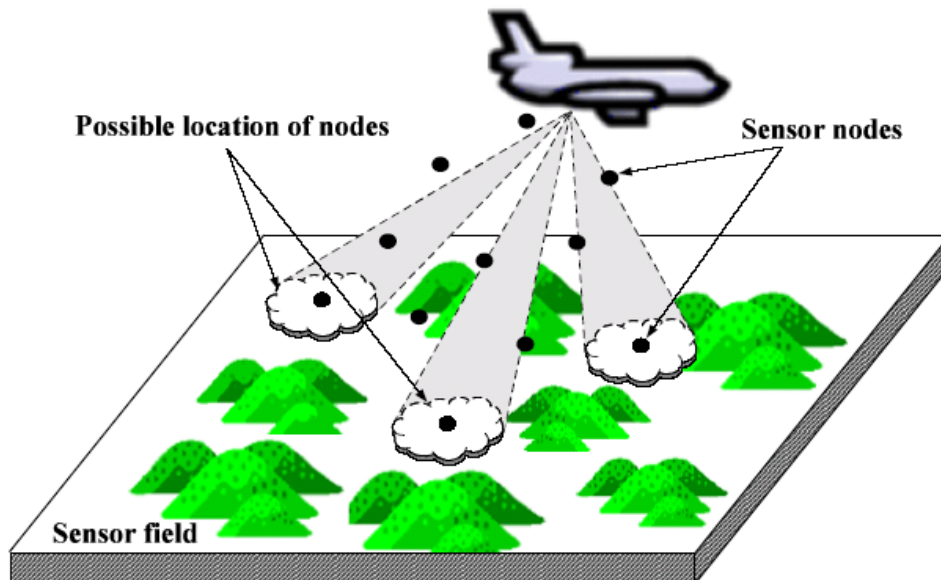
Tab.1: Results of the Convergence Metric for the KUR Test Function

Diversity Metric	NSGA-II	MOPSO
Best	0.3344	0.4803
Worst	0.7825	0.6413
Mean	0.4399	0.5602

Tab.2: Results of the Diversity Metric for the KUR Test Function

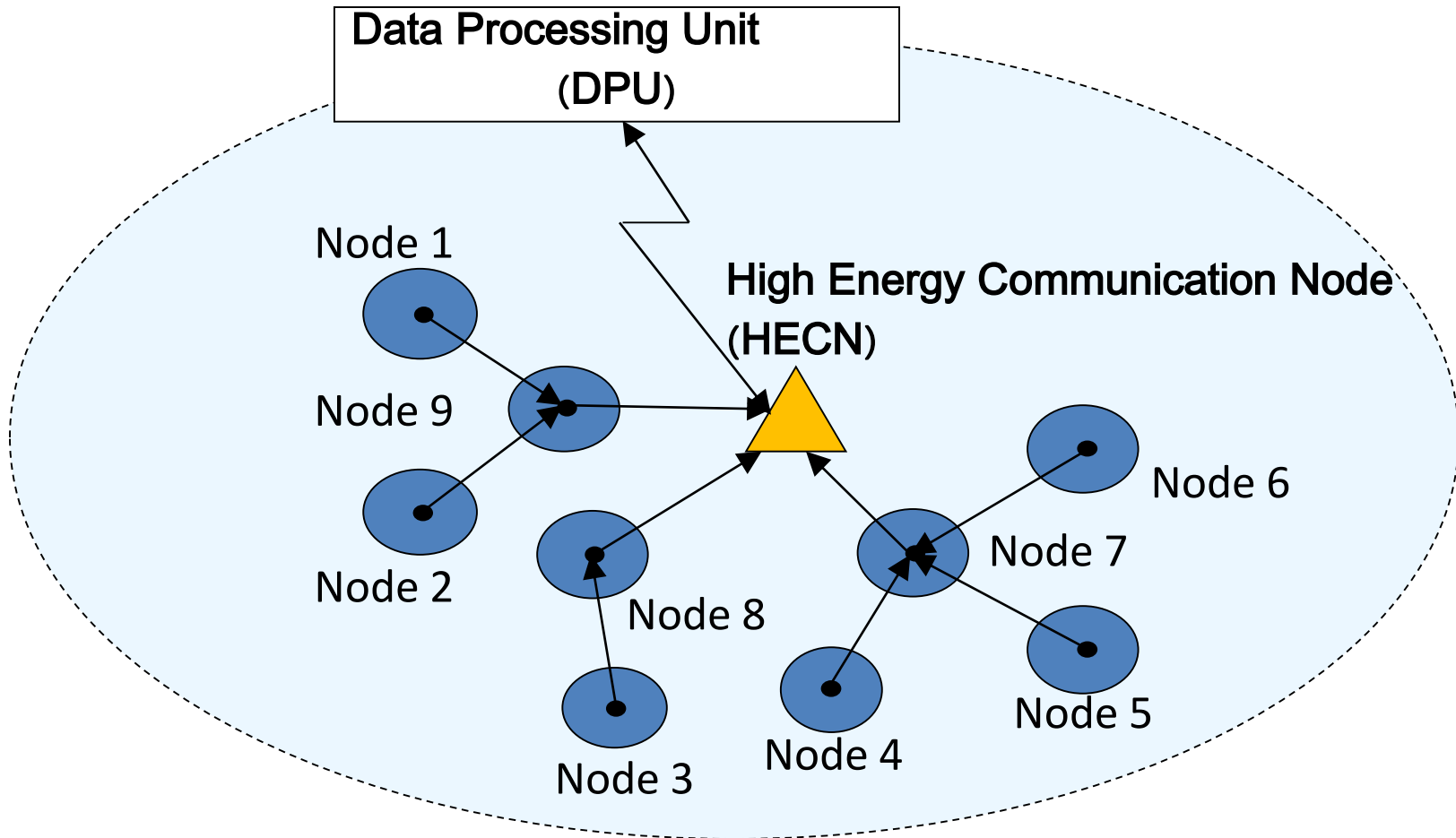
* Total number of fitness evaluations was set to 20,000.

Layout Optimization for a Wireless Sensor Network using NSGA - II



- a) Coverage
- b) Lifetime

Wireless Sensor Network (WSN)



Example of a WSN where sensor nodes are communicating with the DPU through HECN

Optimization of Coverage

- Coverage is defined as the ratio of the union of areas covered by each node and the area of the entire ROI.

$$C = \frac{\bigcup_{i=1, \dots, N} A_i}{A}$$

A_i - Area covered by the i^{th} node

N - Total number of nodes

A - Area of the ROI

Optimization of Lifetime

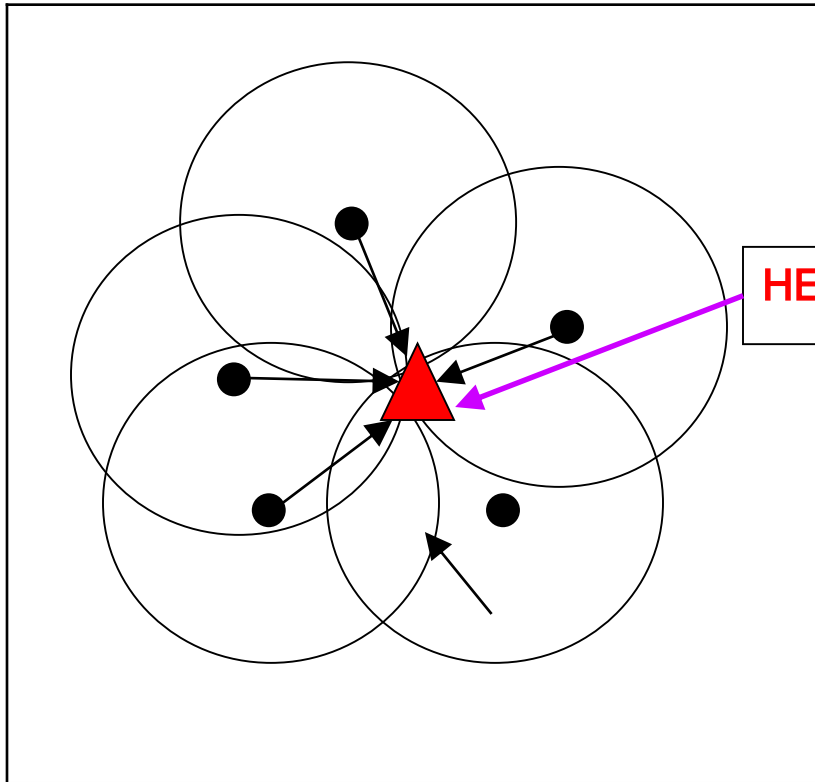
- The lifetime of the whole network is the time until one of the participating nodes run out of energy.
- In every sensing cycle, the data from every node is routed to HECN through a route of minimum weight

$$\text{Lifetime} = \frac{T_{\text{failure}}}{T_{\text{max}}}$$

T_{failure} = maximum number of sensing cycles before failure of any node
 T_{max} = maximum number of possible sensing cycles

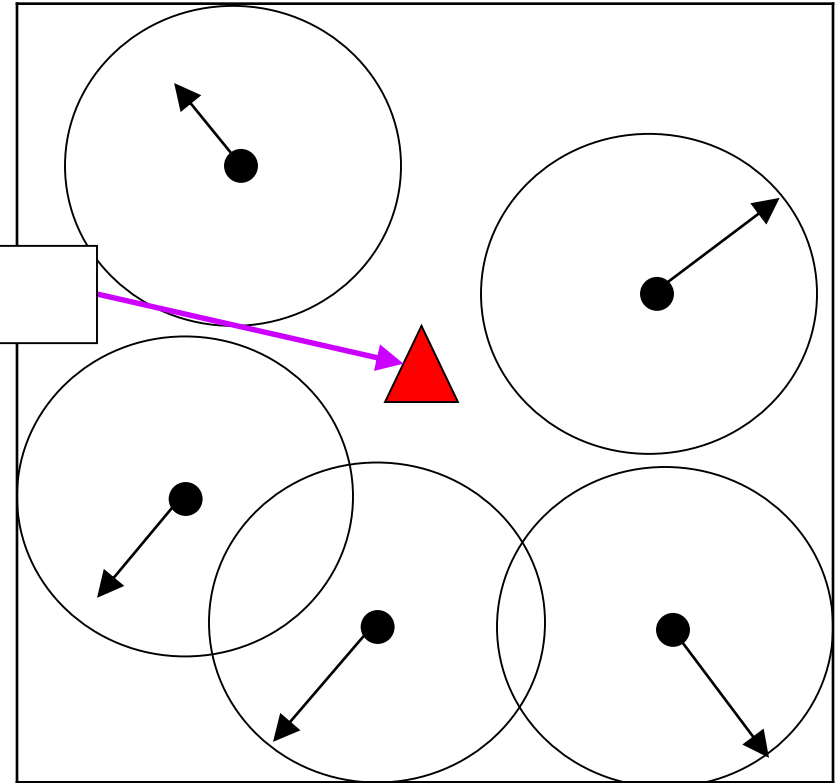
Competing Objectives

Lifetime



- try to arrange the nodes as close as possible to the HECN for maximizing lifetime

Coverage



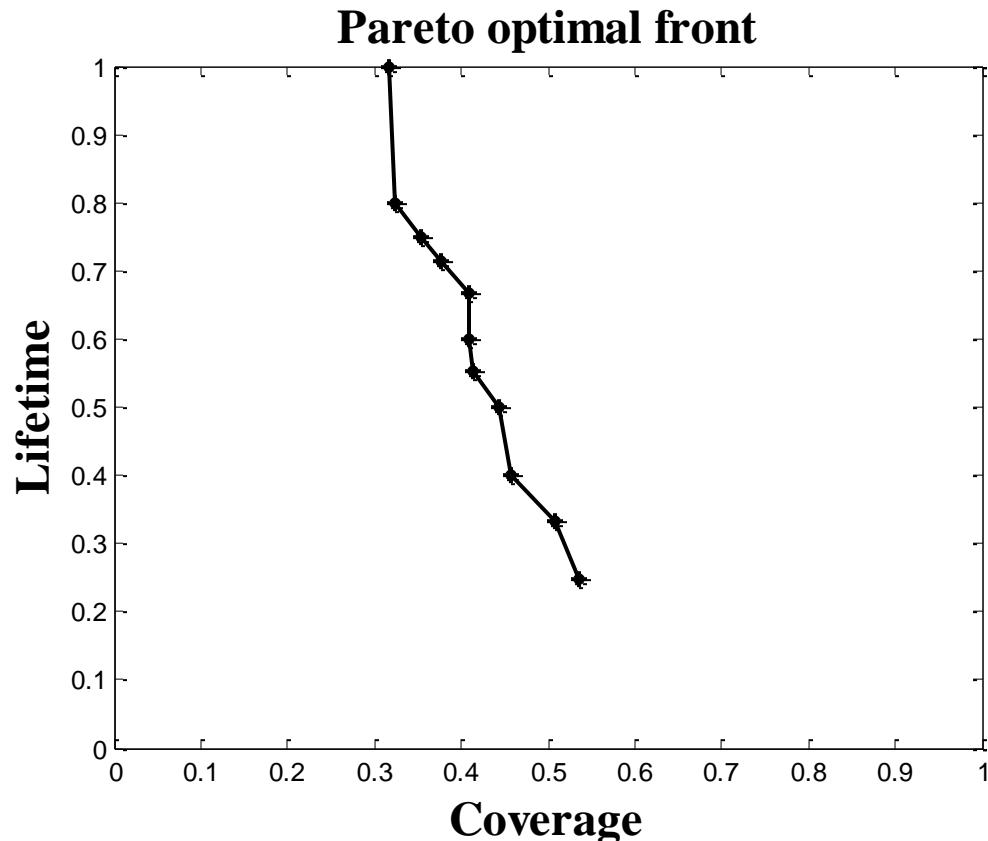
- try to spread out the nodes for maximizing coverage

Simulation Parameters

Parameters of NSGA-II

Number of chromosomes	100
Number of generations	50
Crossover Probability	0.9
Mutation Probability	0.5
Distribution index for crossover	20
Distribution index for mutation	20
Tour size	2

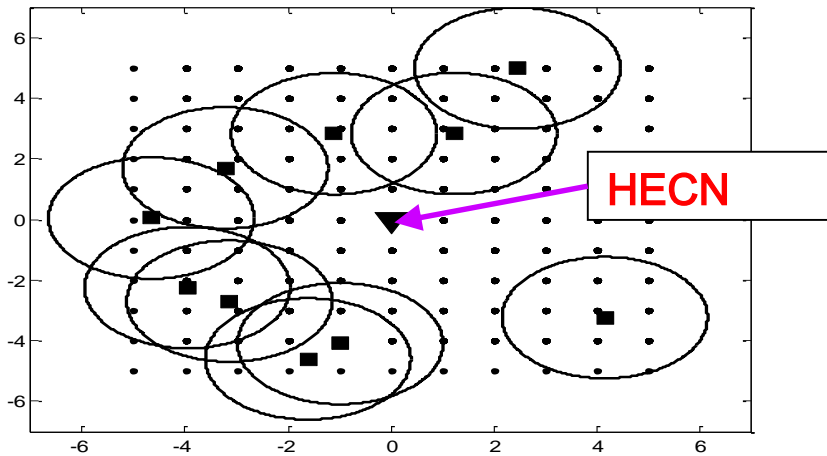
NSGA-II Results



- Pareto Front obtained for a WSN with 10 sensors, 100 chromosomes and 50 generations

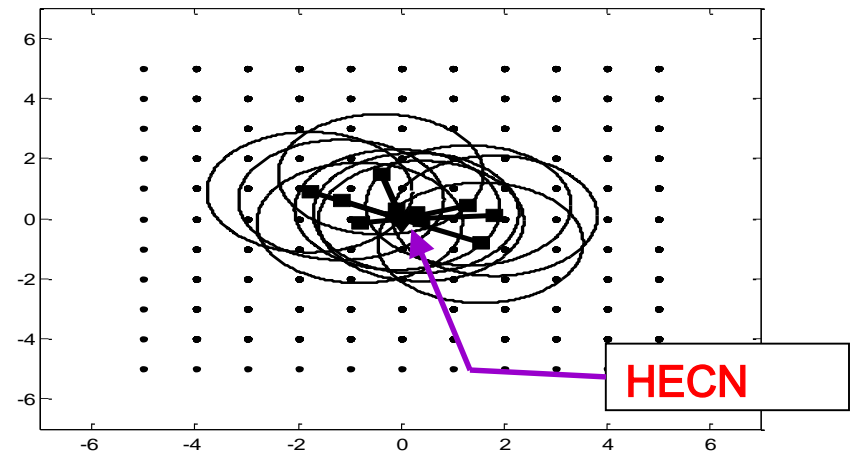
NSGA-II Results (Cont'd)

Initial Disconnect Network



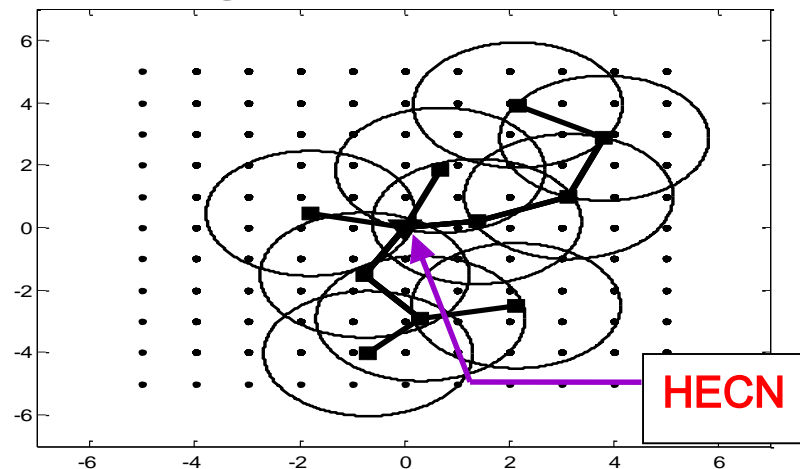
Best Lifetime

Coverage = 0.3335 Lifetime = 0.999



Best Coverage

Coverage = 0.5353 Lifetime = 0.249



References

1. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “***A fast and elitist multi-objective genetic algorithm: NSGA-II***”, *IEEE Transaction on Evolutionary Computation*, 6(2), 181-197, 2002.
2. K. Deb and R. B. Agrawal, “***Simulated binary crossover for continuous search space***,” in *Complex Syst.*, vol. 9, pp. 115–148., Apr. 1995.
3. N. Srinivas and K. Deb, “***Multiobjective function optimization using nondominated sorting genetic algorithms***,” *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, Fall 1995.
4. J. Horn, N. Nafploitis, and D. E. Goldberg, “***A niched Pareto genetic algorithm for multiobjective optimization***,” in *Proceedings of the First IEEE Conference on Evolutionary Computation*, Z. Michalewicz, Ed. Piscataway, NJ: IEEE Press, pp. 82–87 , 1994.
5. J. D. Knowles and D.W. Corne, “***Approximating the nondominated front using the Pareto archived evolution strategy***,” *Evol. Comput.*, vol. 8, pp. 149–172, 2000.
6. Carlos A. Coello Coello, Member, IEEE, Gregorio Toscano Pulido, and Maximino Salazar Lechuga, “***Handling multiple objectives with particle swarm optimization***”, *Evol. Comput.*, vol. 8, pp. 256–279, No. 3, June 2004

THANK YOU