An Intelligent WBAN System for Heart Disease Prediction Using Non-Dominated Genetic Algorithm (ITS-WBAN)

Babak Emami-Abarghouei

Saeed Samet

Presenter ：Ahora Sadeghi-Boroujerdi

Department of Computer Science, (E-health Research Unit), Memorial University of Newfoundland, St. John’s, NL, Canada

E-Health Conference 2016

Jun 5 - 8, 2016 | Vancouver, BC

JUN 2016

Outline

- Introduction (Tele monitoring System)
  - WBAN
- Background:
  - WBAN Application
  - Data Processing Methods for WBAN
    - Optimization
- Problem statement
- Methodology:
  - NSGA Methodology
    - Cost Position (Patient's position)
- Contribution
- Experimental Results

Introduction

Tele Monitoring Health Care System:

An economic application for analyzing the patient's health status

-Remote Patient Monitoring (RPM)

- Ease of Use
- Low Cost
- Comprehensive monitoring system
  - Population of elderly people
  - Medical Centre
- Increase the efficient utilization of physician's skills

Wireless Body Area Network (WBAN)

Trade-off between E-Health Applications and Wireless Sensor network

- Wearable sensor: Biomedical sensor + Wireless device
- Autonomous system
- Remote Patient Monitoring

**

Combination of Wearable sensors , Mobile Computing Center (MCC), Wireless Communication and Medical center

**

WBAN Application

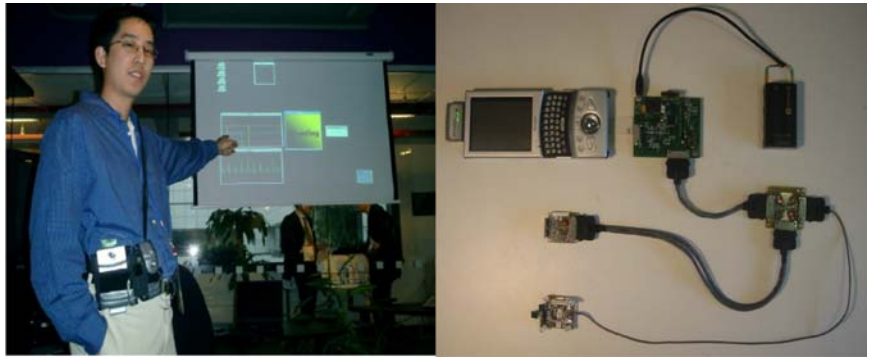
Monitoring System

- Data mining and detection system
  - Time complexity, using machine learning algorithms
    - Offline monitoring (time complexity)

- Prediction system : needs to have a real-time data processing

Data analyzing Methods

- **Machine Learning Algorithm:**
  - Live-net
  - Media Laboratory of MIT, and Cambridge
  - Embedded machine learning –Bayesian



- **Optimization Method:**
  - Classification of two-class of cancer
  - Kanpur Genetic Algorithms Laboratory (KGAL) in India
  - Using NSGA-II to detect the patient lifetime and detect effective genomes

Data Processing Methods using MOA

(Multi Objective Algorithm)

Sorting and Processing Data (Features)

- Multi Objective Problem(MOP)
  - Using mathematical equation to discover all satisfiable condition
- Using Evolutionary Algorithm (MOEA)
  - Multi-Target in one simulation using Pareto Frontier optimization theory

Optimization

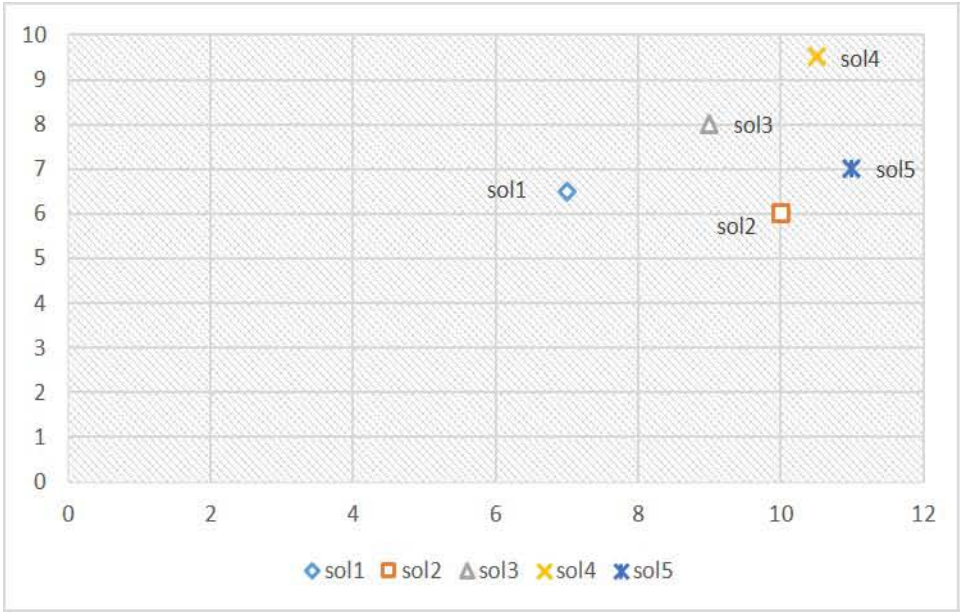
- Using Pareto Frontier (PF) theory To formulate Multi-Target Tracking

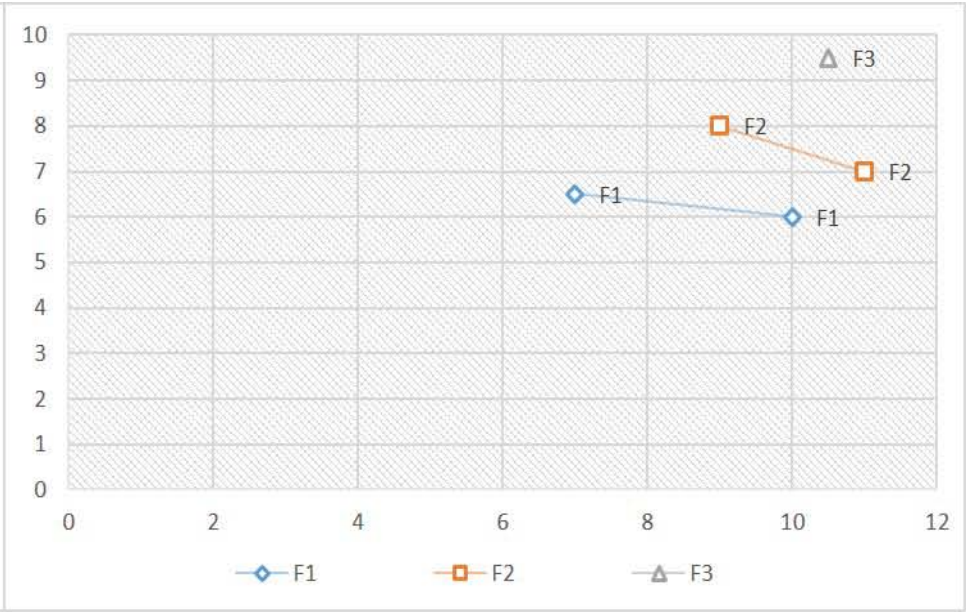
- **PF Terminology:**
  - Trade-off between multiple objects rather than a single Optimum solution (Non-Dominated Solutions)
  - Comparing all features of each solution with all features of other solutions - elitism selection (non-dominated set)

- **NSGA**:
  - a feasible optimization method using genetic algorithm

Using NSGA-II to Sort Data based on Dominated Rank

	Objective-1	Objective-2	Dominated rank, dominated solutions	
sol1	7	6.5	0	{3,4,5}
sol2	10	6	0	{3,4,5}
sol3	9	8	2	{5}
sol4	10.5	9.5	3	{}
sol5	11	7	2	{5}





NSGA Methodology

This method has been designed to visualize data in 2-Dimension graph (patient's position)

- Cost Function: trade-off between :
  - Static data :Age; Chest Pain; Cholesterol; Fasting Blood Sugar; Resting ECG; Exercise Induced Angina; Old peak; Number of Colored Vessels
  - Dynamic Data :Resting Blood Pressure ; Normal Heartbeat Rate

Contribution

- While using classification methods only determine whether or not the patient is in healthy case, by applying an optimization approach the critical point for each patient can be explored in linear time as well.

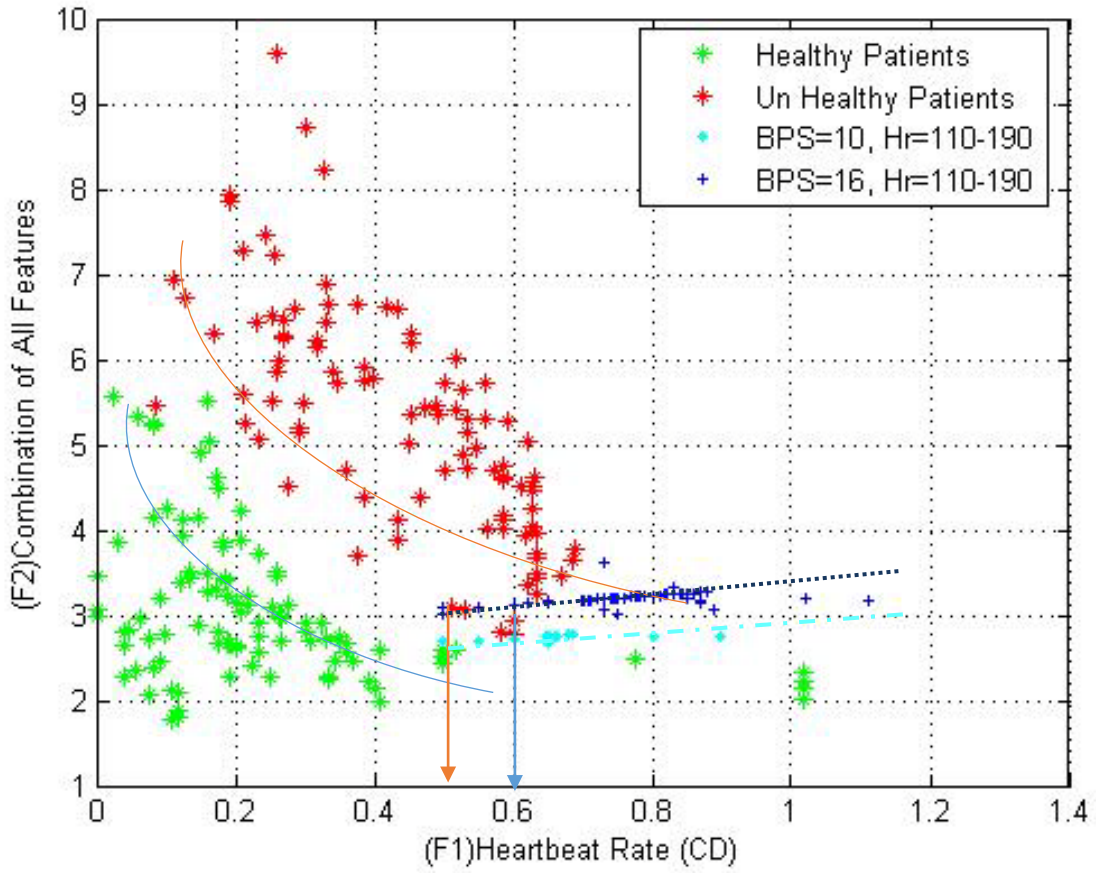
- We are taking advantage using NSGA-II to determine the likelihood of closeness of a patient, which is categorized in a specific class, to the border of the other class.

- Next we explored how phenotype would be changed while some dynamic variable are changed (Heartbeat rate, Blood pressure)

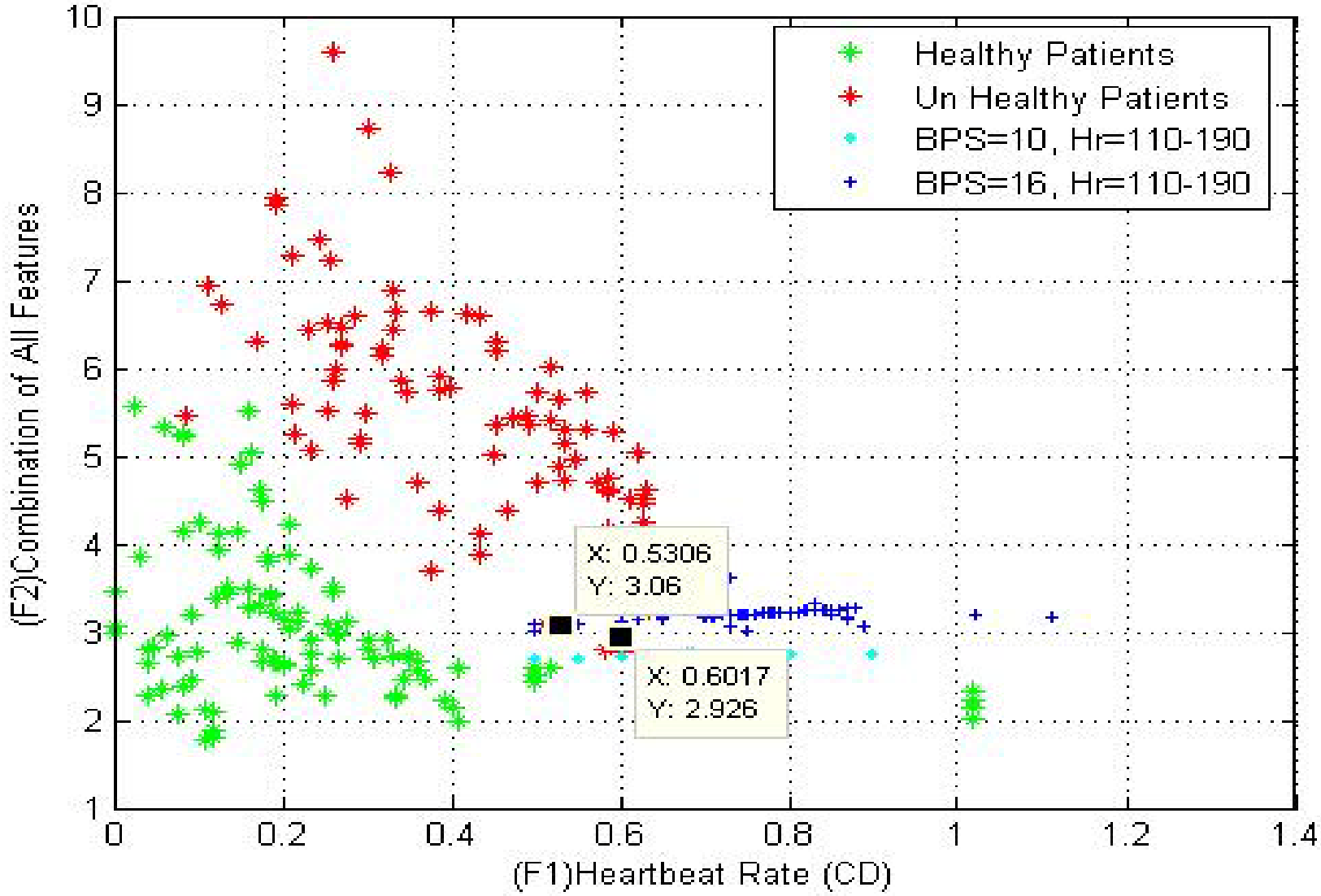
Simulation Setup

Simulator software	Matlab 2011
Dataset	Heart disease Dataset (Cleveland)
Normal cases	150
Abnormal cases	120
Mutation ratio	0.2
Parent selection	binary tournament
Number of generation	50
Selected population	100

PF curve after 50 generation



HR	Age	(cp)	rest bps	col	chol	fbs	restecg	exang	oldpeak	oldpeak	slope	ca	Class
0.5	**0.1**	**0.33**	**0.5**	**0.3**	**0.25**	**0**	**0.5**	**0**	**0**	**0**	**0**	**0**	0
0.65	**0.1**	**0.33**	**0.5**	**0.3**	**0.25**	**0**	**0.5**	**0**	**0**	**0**	**0**	**0**	1
0.55	**0.1**	**0.33**	**0.7**	**0.3**	**0.25**	**0**	**0.5**	**0**	**0**	**0**	**0**	**0**	1
HR 60=0, 0.5 =95, 0.53=110, 0.60=130													
BPS, 8= 0, 0.5 = 12, 0.7 =15													


Healthy Patients
Un Healthy Patients
BPS=10, Hr=110-190
BPS=16, Hr=110-190
X: 0.5306
Y: 3.06
X: 0.6017
Y: 2.926
(F2)Combination of All Features
(F1)Heartbeat Rate (CD)
10
9
8
7
6
5
4
3
2
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4

Conclusion

- In this study, a threshold of healthy and unhealthy phenotypes are extracted using NSGA-II.

- In future, we are going to have a real time prediction for our patients due to their activities. (Introducing the new WBAN application named Self Coach as a real-time monitoring system)

Selected References

- [1] C. A. C. Coello, D. A. Van Veldhuizen, and G. B. Lamont, *Evolutionary algorithms for solving multi-objective problems* vol. 242: Springer, 2002.
- [2] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," *Lecture notes in computer science*, vol. 1917, pp. 849-858, 2000.
- [3] Deb, K. and A.R. Reddy, Reliable classication of two-class cancer data
- using evolutionary algorithms. BioSystems, 2003. 72(1): p. 111-129.
- [4] Michael Sung, C.M.a.A.P., Wearable feedback systems for rehabilitation. Journal of NeuroEngineering and Rehabilitation, 2005. 2(17): p. 12.

Thank You !

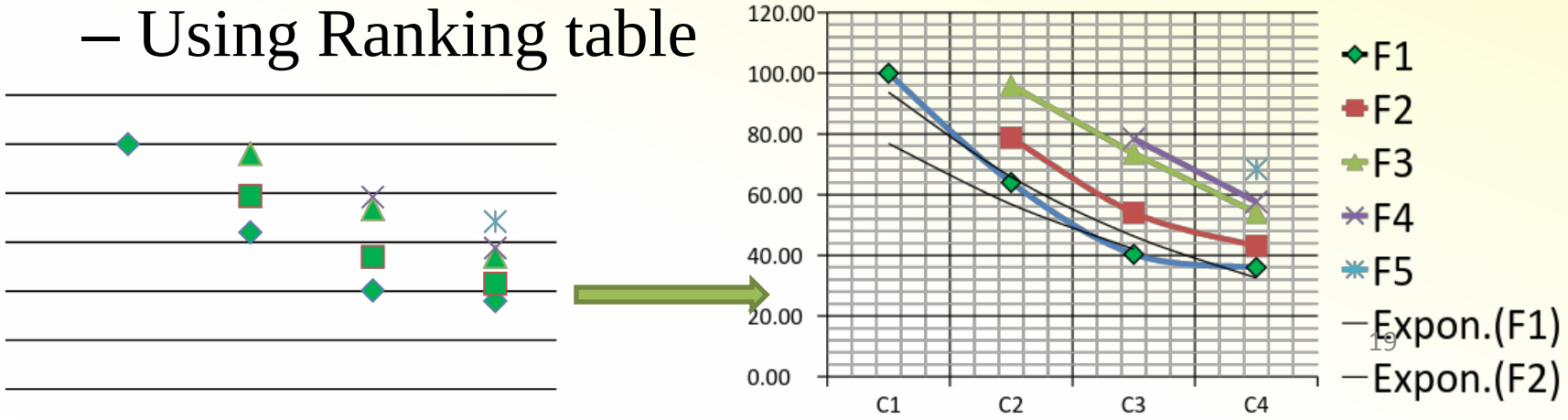
Question and Comments

Problem statement

- Using common method of classification has time complexity, not efficient for the real-time data processing.
- Some features have been omitted using feature selection such as age, and patient's activity, due to speed up the classification
- Using Classification method cannot determine the risk percentage of each patient either in normal or disease class (the likelihood of closeness of a patient, which is categorized in a specific class, to the border of the other class.)

Sorting Data (Pareto Front)

- Data Sorting based on dominated level
  - Non-Dominated Data
  - first step of sorting
  - Using Ranking table



Appendix-Ranking table for Domination

	G1	G2
sol1	7	6.5
sol2	10	6
sol3	9	8
sol4	10.5	9.5
sol5	11	7

Dominate matrix

	sol1	sol2	sol3	sol4	sol5
sol1	**0**	**0**	**1**	**1**	**1**
sol2	**0**	**0**	**0**	**0**	**1**
sol3	**0**	**0**	**0**	**1**	**0**
sol4	**0**	**0**	**0**	**0**	**0**
sol5	**0**	**0**	**0**	**0**	**0**

	sol1	sol2	sol3	sol4	sol5
sol1	0	0	1	1	1
sol2	0	0	0	0	1
sol3	0	0	0	1	0
sol4	0	0	0	0	0
sol5	0	0	0	0	0
	0	0	1	2	2

F1={1 ,2}

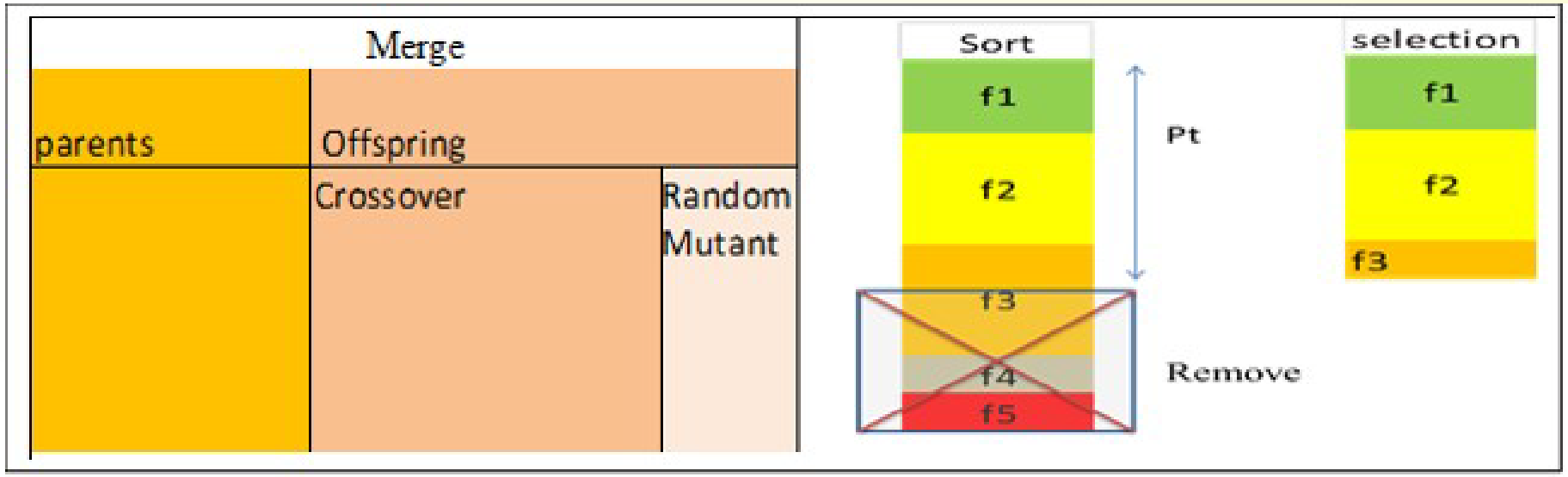
	sol1	sol2	sol3	sol4	sol5
sol1	inf	inf	0	0	0
sol2	inf	inf	0	0	0
sol3	inf	inf	0	1	0
sol4	inf	inf	0	0	0
sol5	inf	inf	0	0	0
	inf	inf	0	1	0

F2={3 ,5}

	sol1	sol2	sol3	sol4	sol5
sol1	inf	inf	inf	0	inf
sol2	inf	inf	inf	0	inf
sol3	inf	inf	inf	0	inf
sol4	inf	inf	inf	0	inf
sol5	inf	inf	inf	0	inf
sum	inf	inf	inf	0	inf

F3={4 }

- Merge population (Crossover , Nutation , Generation)
- Sort population
- Tournament Selection



Crowding Distance (CD)

$$CD_i^{\,1} = \frac{|f_1^{\,i+1} - f_1^{\,i-1}|}{f_1^{\,Max} - f_1^{\,Min}} \quad , CD_i^{\,2} = \frac{|f_2^{\,i+1} - f_2^{\,i-1}|}{f_2^{\,Max} - f_2^{\,Min}}$$

$$D_i = CD_i^{\,2} + CD_i^{\,1}$$