

Original Research Article

Genetic Algorithm based on K-means-Clustering Technique for Multi-objective Resource Allocation Problems

ABSTRACT

This paper presents genetic algorithm based on K-means clustering technique for solving multi-objective resource allocation problem (MORAP). By using k-means clustering technique, population can be divided into a specific number of subpopulations with dynamic size. In this way, different GA operators (crossover and mutation) can be applied to each subpopulation instead of one GA operators applied to the whole population. The purpose of implementing K-means clustering technique is preserving and introducing diversity. Also it enable the algorithm to avoid local minima by preventing the population of chromosomes from becoming too similar to each other. Two test problems taken from the literature are used to compare the performance of the proposed approach with the competing algorithms. The results have been demonstrated the superiority of the proposed algorithm and its capability to solve MORAP.

Keywords: Multi-objective resource allocation problems, Genetic algorithm, K-means clustering technique, optimization.

1. INTRODUCTION

Resource allocation is part of resource management. It is used to assign the available resources in an economic way. In project management, resource allocation is the scheduling of activities and the resources required by those activities while taking into consideration both the resource availability and the project time [1]. MORAP is the process of allocating resources among the various projects or business units to meet the expected objectives. Resources may be manpower, dominions, raw materials, capital or anything else in limited supply which can be used to accomplish the goals. The goals may be objectives or targets (i.e., maximizing profits, minimizing costs, or achieving the best possible quality) [2]. MORAP has a variety of applications such as:

- 1- Allocating Marketing Resources [3]: A key responsibility of marketing managers is to decide the allocation of scarce marketing resources, e.g., advertising dollars, selling hours, retail shelf-space or merchandise inventories.
- 2- Stochastic network systems [4]: for getting the optimal performance of the network by maintaining both resource and duration of activities in proper realization.
- 3- Scheduling of Worker Allocation in the Manual Labor environment [5]: for assigning the workers into the jobs to reduce human cost, to shorten production duration and control production overwork.
- 4- Portfolio optimization [6]: for creating efficient portfolios on allocating funds to stocks or bonds to maximize return for a given level of risk, or to minimize risk for a target rate of return.
- 5- Health care resource allocation [7]: Discounting costs and health benefits in cost-effectiveness analysis for the health care resource allocation. And many other application that can be formulated as resource allocation problem [8].

41
42 Traditionally, MORAP have been solved using methods in operations research (OR), Integer
43 Programming (ID) [9], branch-and-bound (B&B) [10], and Dynamic Programming (DP) [11].
44 None of these methods are computationally tractable for any real-life problem size, thus
45 rendering them impractical [12].

46 In recent years, there has been a growing interest in studying evolutionary algorithms (EAs)
47 for many real world optimization problems, which are conceptually different from the
48 traditional mathematical programming techniques. These methods are based on certain
49 biological, molecular, and neurological phenomena. The reasons for their popularity are
50 many [13]:

- 51
52 (i) EAs do not require any derivative information
53 (ii) EAs are relatively simple to implement
54 (iii) EAs are flexible and have a wide-spread applicability
55

56 One of the main advantages of these EAS is finding high quality or near-optimal solutions
57 with reasonable computational times. Recently, several evolutionary algorithms have been
58 developed for MORAP including genetic algorithm (GA) [14], genetic algorithm with fuzzy
59 inference [15], A hybrid simulated annealing approach [16], ant colony optimization (ACO)
60 [2], variable neighborhood search [17], a memetic algorithm based on node-weighted graphs
61 [18], hybrid particle swarm optimization [19], tabu search [20], differential evolutionary
62 algorithm [21], multi-objective differential evolution based on weighted normalized sum
63 (WNS-MODE) approach [22], co-evolutionary hyper-heuristic method [23], combining
64 exhaustive search with evolutionary computation method [24], Pareto multistage decision-
65 based genetic algorithm (P-mdGA) [25], an improved quantum evolutionary algorithm (QEA)
66 [26].

67 Clustering is a process of division of data into groups of similar objects. Each group, called
68 cluster, consists of objects that are similar between themselves and dissimilar to objects of
69 other groups [27]. There are many clustering algorithms [28]. The k-means is possibly the
70 most commonly-used clustering algorithm because of its simplicity and accuracy. The k-
71 means finds a locally optimal solution by minimizing a distance measure between each data
72 and its nearest cluster center [29]. Several algorithms have been proposed in the literature
73 for clustering: The Iterative Self-Organizing Data Analysis Technique (ISODATA) [30],
74 Clustering Large Applications based up on Randomized Search (CLARANS) [31], Parallel-
75 cluster (p-cluster) [32], Density-Based Spatial Clustering of Applications with Noise
76 (DBSCAN) [33] and Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH)
77 [34]. The most commonly and widely used clustering is K-Means because of its simplicity
78 and accuracy [29].
79

80 In particular, genetic algorithms (GA) produce good results compared to other techniques in
81 solving resource allocation problem [35]. The major disadvantage of GA, though, is that they
82 easily become trapped in the local minima. In this paper, k-means clustering technique is
83 used along with a GA in order to avoid the local minima problem and to benefit from the
84 advantages of dynamic clustering algorithms. By using K-means clustering algorithm,
85 population can be divided into a specific number of subpopulations. Within each cluster,
86 subpopulation has common features. After grouping all individuals into pre-defined number
87 of clusters, instead of taking care of all individuals, population can regard the huge amount
88 of individuals as just the number of groups that has been divided. In this way, different GA
89 operators can apply to subpopulations instead of one GA operator applied to all population.
90

This paper is organized as follows. In section 2, we describe the mathematical formulation of the MORAP. In sections 3, genetic algorithm is briefly introduced. In section 4, clustering technique is briefly introduced. **Our approach** is proposed and explained in detail in section 5. Numerical simulation is presented in section 6. Finally, we conclude the paper in section 7.

2. MATHEMATICAL FORMULATION

The general form of the MORAP is as follows [14]:

$$\begin{aligned}
 \text{Max} \quad & Z_1(x_1, x_2, \dots, x_n) = \sum_{k=1}^n z_k^1(x_k) \\
 \text{Max} \quad & Z_2(x_1, x_2, \dots, x_n) = \sum_{k=1}^n z_k^2(x_k) \\
 & \dots\dots\dots \\
 \text{Max} \quad & Z_q(x_1, x_2, \dots, x_n) = \sum_{k=1}^n z_k^q(x_k) \tag{1} \\
 \text{Subject to} \quad & \sum_{k=1}^n g_k(x_k) \leq S, \\
 & g_k(x_k) \geq 0, \\
 & x_k \geq 0;
 \end{aligned}$$

where x_k define (decision variables), S donates the (resources), $g_k(x_k)$ represents the activities stages and $Z_q(x_1, x_2, \dots, x_n)$ is the q objective function. The process of MORAP seek to find an optimal allocation (decision variables) of limited amount of resource to a number of tasks (activities stages) for optimizing their objective subjected to the given resource constraint [14].

3. GA PROCEDURE

GAs operates on a population of candidate solutions encoded to finite bit string called chromosome. In order to acquire optimality, each chromosome exchanges information by using operators borrowed from natural genetic to produce the improved solution [36, 37].

We can describe the steps of basic genetic operators as follow [38]:

- Step 0 (Initialization stage): Randomly generate an initial population.
- Step 1 (Evaluation of non-dominated solutions): Calculate the values of q objective function for each individual.
- Step 2 (Selection operators): Selection operators are used to select the individuals to which the crossover operators will be applied.
- Step 3 (Crossover operators): In this step, crossover operators are applied to each pair selected to generate offspring
- Step 4 (Mutation operators): Apply mutation operation to each offspring generated by the crossover operation.
- Step 5 (Stopping criterion): check the stopping criterion, if it is not reached, then go to Step 1, otherwise, go to the next step.
- Step 6 (DM selection): DM selects the most preferred among the alternatives.

The basic step of GA is illustrated in flowchart as shown in figure 1

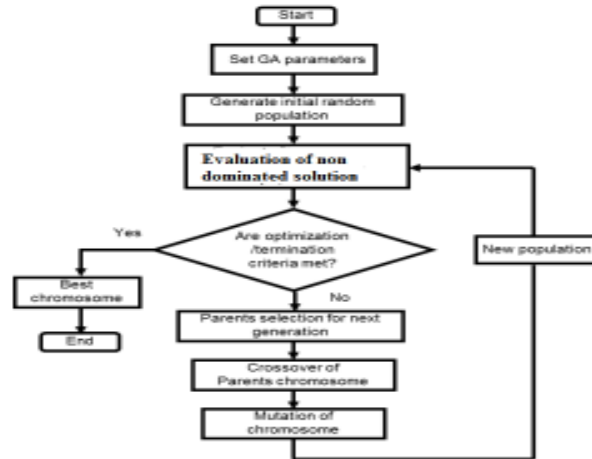


Fig.1. Main flowchart of GA (Taken from [39]).

4. CLUSTERING ALGORITHM

Clustering is process of Finding groups of objects such that the objects in a group will be similar (or related) to each other and different from (or unrelated to) the objects in other groups [27]. K-means [29] is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group is done. At this point we need to re-calculate k new centroids as centers of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. The loop has been generated, as a result of this loop we may notice that the k centroids change their location step by step until no more changes are done. In other words centroids do not move any more. A centroid is defined as the point whose coordinates are obtained by computing the average of each of the coordinates (i.e., feature values) of the points of the cluster. Formally, the k-means clustering algorithm follows the following steps (taken from [40]):

Step 1: define a number of desired clusters, k .

Step 2: choose initial cluster centroid randomly, These represent the “temporary” means of the clusters.

Step 3: compute the squared Euclidean distance (sum of square error) from each object to each cluster and each object is assigned to the closest cluster as follows:

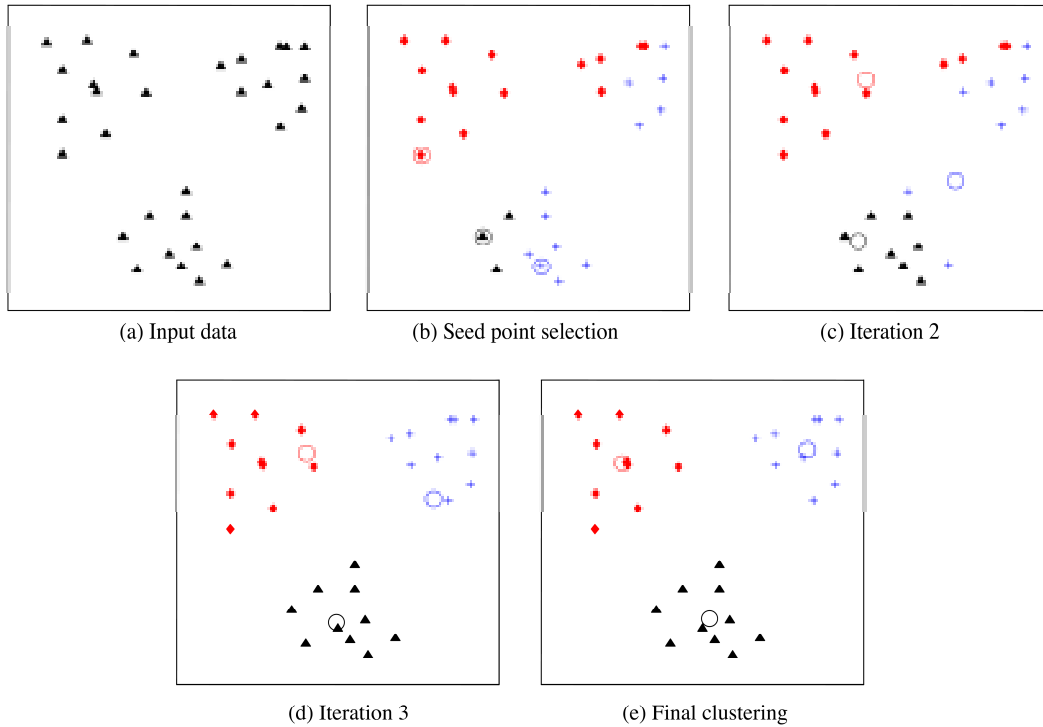
$$\text{Sum of Squire Error (SSE)} = \sum_{i=1}^k \sum_{x \in C_i} \text{dist}^2(m_i, x); \quad (2)$$

where x is a data point in cluster C_i and m_i is the centroid of cluster C .

Step 4: For each cluster, the new centroid is computed, and each centroid value is now replaced by the respective cluster centroid.

Step 5: Repeat steps 3 and 4 until no point changes its cluster.

167 Figure 2 shows an illustration of K-means algorithm on a 2-dimensional dataset with three
 168 clusters.
 169



170
 171 **Fig. 2. Illustration of K-means algorithm. (a) Two-dimensional input data with three**
 172 **clusters; (b) three centroid points selected as cluster centers and initial assignment of**
 173 **the data points to clusters; (c) & (d) intermediate iterations updating cluster labels**
 174 **and their centers; (e) final clustering obtained by K-means algorithm at convergence.**
 175 **(Taken from [41])**
 176
 177

178 5. GENETIC ALGORITHM BASED ON K-MEANS-CLUSTERING TECHNIQUE

179
 180 In this paper, we propose GA based on k-means clustering technique to solve MORAP.
 181 Where, population can be divided into a specific number of groups (subpopulations) by
 182 clustering technique. Within each cluster, subpopulation has common features. After
 183 grouping all individuals into pre-defined number of clusters, instead of taking care of all
 184 individuals, population can regard the huge amount of individuals as just the number of
 185 groups that has been divided. In this way, different GA operators can applied to sub
 186 population instead of one GAs operator applied to all population. The main steps of the
 187 proposed algorithm are described as follows:
 188

189 5.1 Initialization stage

190
 191 In this paper, we reformulate MORAP as network model, where limited supply represented
 192 by stages as shown in figure 3. We consider the human resource allocation problem with
 193 multi-objectives for minimize the total cost and maximize the total efficiencies of the human
 194 allocation decision. It is required to find a path between two nodes source node (S) and
 195 terminal node (T) having minimum total cost and maximum efficiency [14]. A path from node

196 S to node T is a sequence of arcs $(S, x_{1m}), (x_{1m}, x_{2m}), \dots (x_{n-1m}, T)$. A path can be
 197 equivalently represented as sequence of nodes $(S, x_{1m}, x_{2m}, \dots, x_{n-1m}, T)$; where
 198 $m = 0, 1, 2, \dots, m$ and n is number of stage (district), thus the structure of each individual
 199 (chromosome) can express a path, i.e., each chromosome is represented as sequence of
 200 nodes.

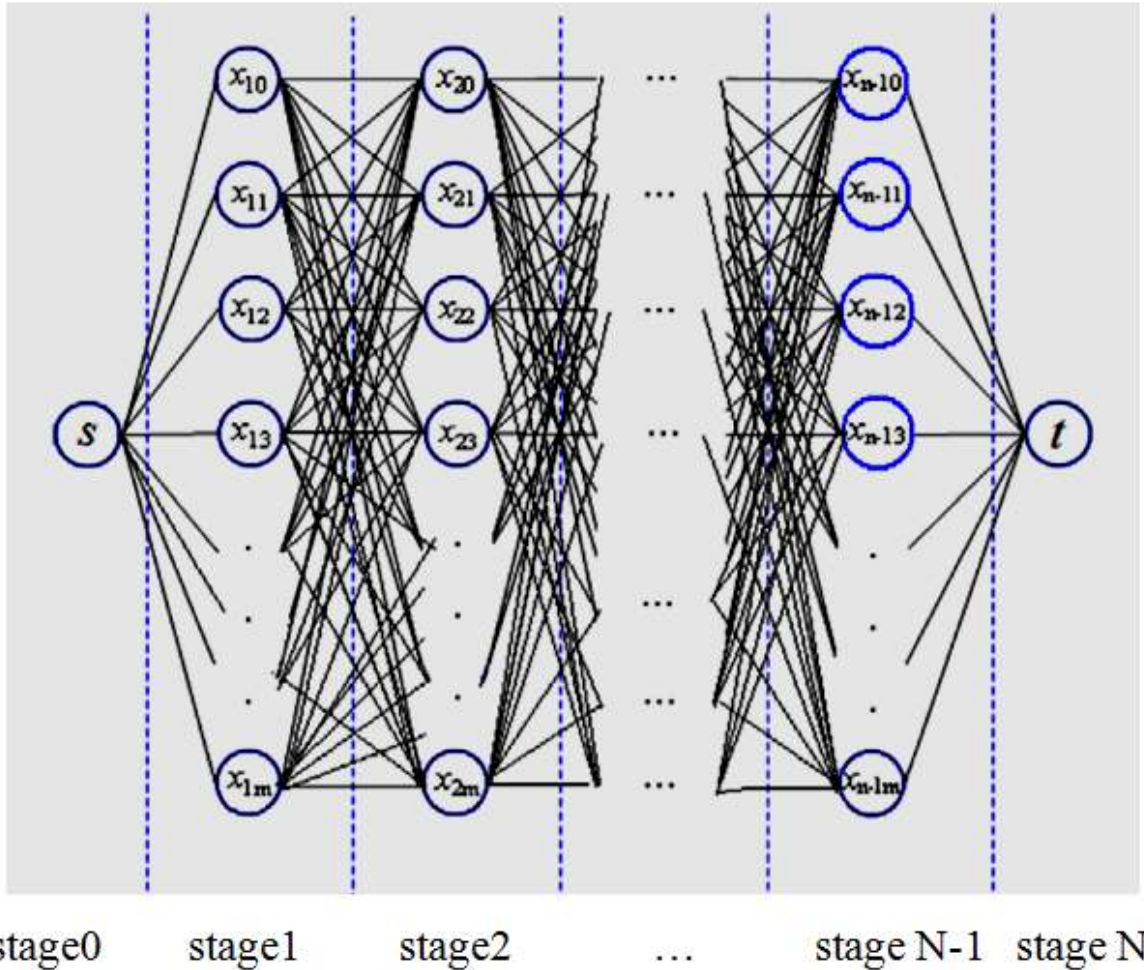


Fig. 3. Representation MORAP as network model

In this step, the algorithm generates an initial population containing $Npop$ strings. We use a random selection element from available number in each stage. For example, in the MORAP of allocating 10 worker to a certain set of 4 jobs (4 stages), the multi-objective resource allocation considers determining a vector of allocation path in the 11 states and the 4 stages under the minimum costs and the maximum efficiency as shown in figure 4. Figure 5 shows the structure of the chromosome for four stage allocation path.

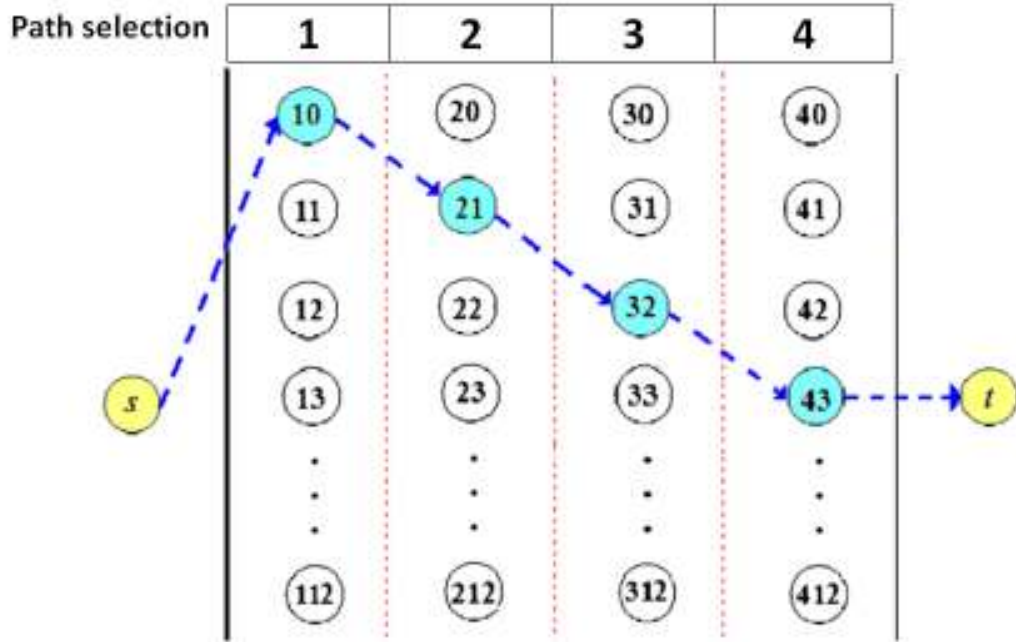


Fig. 4.The figure of optimal path allocation

Stage :	1	2	3	4
Path Selection	3	2	5	0

Fig. 5.The structure of the chromosome for four-stage allocation path

5.2 Rejection of illegal individuals

The chromosome structure may produce some illegal individuals, which do not satisfy the feasible region of the problem, for example, if there are N workers needed to allocate to M jobs, the number of worker must not exceed N workers [14]. So before the genetic operation for every time, we should eliminate these illegal individuals which do not satisfy the conditions mentioned above. In this step, the illegal individuals are simply eliminated and replaced by randomly drawn new legal individuals.

5.3 Evaluation of non-dominated solutions

In multi-objective optimization problem, there does not exist a single solution that simultaneously optimizes each objective. In that case, the objective functions are said to be discordant and there exists a (possibly infinite) number of solutions. For such solutions, called non-dominated solutions, Pareto optimal solutions, Pareto efficient solutions or noninferior solutions [42]. So it is necessary to find the Pareto optimal solutions by classifying a population according to non-domination.

Definition (Pareto optimal solution): x^* is said to be a Pareto optimal solution of multi-objective optimization problem if there exists no other feasible x (i.e., $x \in S$) such that, $f_j(x) \leq f_j(x^*)$ for all $j = 1, 2, \dots, q$ and $f_j(x) < f_j(x^*)$ for at least one objective function f_j .

237 The image of the Pareto set, i.e., the image of all the Pareto solutions, is called Pareto front
 238 or Pareto curve or surface. A population can be evaluated according to non-domination
 239 criteria. Consider any two solutions x_1 and x_2 for a problem having more than one objective
 240 function $q(q > 1)$ objective function values, may have one of these possibilities, one of
 241 these solution can be dominates the other or non-dominates the other. The solution x_1 is
 242 said to dominate other solution x_2 if the two following condition are true [43].

243

- 244 1- The solution x_1 is no worse (say the operator ($\prec$) denotes worse and ($\succ$) denotes
 245 better) than x_2 in all objectives, or $f_k(x_1) \prec f_k(x_2)$ for all $k = 1, \dots, q$ objectives.
- 246 2- The solution x_1 is strictly better than x_2 in at least one objective, or $f_k(x_1) \succ f_k(x_2)$ for at
 247 least one $k \in \{1, 2, \dots, q\}$.

248

249 Consider a set of population members, each having $q(q > 1)$ objective function values. The
 250 following procedure is used to find the non-dominated set of solutions [44].

251

252 Step 0: Begin with $m = 1$.

253 Step 1: For all $n = 1, 2, \dots, N_{POP}$ and $m \neq n$, compare solutions x_m and x_n for domination
 254 using the previous conditions for all q objectives.

255 Step 2: If for any n , x_m is dominated by x_n , mark x_m as 'dominated', and it is inefficient.

256 Step 3: If all solutions (that is, when $m = N_{POP}$ is reached) in the set are considered, Go to
 257 Step 4, else increment m by one and Go to Step 1.

258 Step 4: All solutions that are not marked 'dominated' are non-dominated solutions.

259

260 All these non-dominated solutions are assumed to constitute the non-dominated front in the
 261 population in a specified generation.

262

263 5.4 Selection stage

264

265 The main purpose of the selection stage (parent selection stage) is to determine which
 266 individuals are best suited to have children and pass their chromosomes to the next
 267 generation according to their fitness. Better the fitness, the bigger chance to be selected to
 268 be the parent [39] Selection (reproduction) operator is intended to improve the average
 269 quality of the population by giving the high quality chromosomes a better chance to get
 270 copied into the next generation. The selection directs GA search towards promising regions
 271 in the search space.

272

273 Murata et al. [45] proposed a dynamic-weight approach to obtaining fitness function. In
 274 weighted sum approach, multiple objective functions combined into a scalar fitness solution.
 275 The weighed sum objective is given as follows:

$$276 \quad f(x) = w_1 \cdot f_1(x) + \dots + w_i \cdot f_i(x) + \dots + w_q \cdot f_q(x); \quad (3)$$

277 where x is a string (i.e., individual), $f(x)$ is a combined fitness function, $f_i(x)$ is the i th
 278 objective function and w_i is a constant weight for $f_i(x)$, where w is a weighting-vector with
 279 $w_i \geq 0$ for all $i = 1, \dots, q$ and $\sum_{i=1}^q w_i = 1$. In general, the value of each weight can
 280 determine randomly. For a multi-objective optimization problem with q objective functions,
 281 we can assign a random real number to each weight as follows:

$$w_i = \frac{rand_i}{\sum_{j=1}^q rand_j} = 1, \quad i = 1, 2, \dots, q; \quad (4)$$

where $rand_i$ and $rand_j$ are non-negative random real numbers.

In this paper, Binary Tournament selection [46] is used. In the Binary Tournament selection, two individuals are chosen at random and the better of the two individuals is selected and copied in mating pool.

5.5 K-means clustering technique

In order to keep diversity and to avoid trapping in local minima, the k-means cluster algorithm was implemented. In this step, the population in mating pool is split into K separated subpopulations with dynamic size, as illustrated in figure 6.

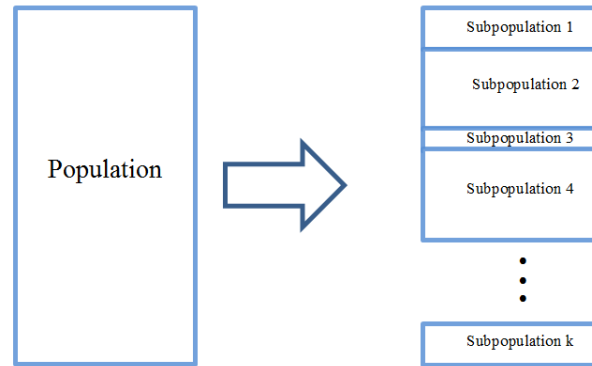


Fig. 6. The population is split into K separated subpopulations with dynamic size

5.6 Genetic algorithms (GAs) operators

GA uses two operators to generate new solutions from existing ones: crossover and mutation. In this step, different GA operators applied to subpopulations instead of one GA operator applied to the whole population. In a simple GA, there is only one population in each generation and all the genetic operations are applied on it, but here, population is split into separated subpopulations, so all the genetic operations, including, crossover and mutation on each subpopulation are performed separately to create new generations.

5.6.1 Crossover operator

The goal of crossover is to exchange information between two chromosomes in order to produce two new offspring for the next population [47]. In our study we used common crossover techniques in a GA. A brief explanation of these techniques is given below.

5.6.1.1 One-point crossover

In one point crossover operator that randomly selects a crossover point within a chromosome then interchanges the two parent chromosomes at this point to produce two new offspring [48].

5.6.1.2 Two-point crossover

322

323 In two-point crossover operators two different cut-off points were randomly selected. New
324 offspring were obtained by relocating zones between the cut-off points of parent
325 chromosome [48].

326

5.6.1.3 Uniform crossover

328

329 In this method, a random binary string is generated with the same-size of chromosome.
330 Then relative genes under this binary string between parents is exchanged, where parent
331 strings exchange their bit at the position where the corresponding position in random binary
332 string is 1. Otherwise, no exchange of bit is performed [49].

333

5.6.1.4 Cross crossover

335

336 In such crossover, the new offspring are selected from different parts of parent; often the
337 produced generation is very different from their parent. A cut point is selected in the middle
338 point in the two parent, the left side gene of the cut point of the first parent are copied directly
339 to right side of the new first offspring. Then, the remaining portion of offspring array is
340 selected from left side gene of the cut point of the second parent. In order to produce the
341 second offspring the parents are swapped in this process as pervious manner [50].

342

5.6.2 Mutation operator

344

345 Mutation operator is one of the GA operators that used to produce new chromosomes or
346 modify some features of them depending on some small probability value. The objective of
347 this operator is to prevent falling of all solutions in population into a local optimum of solved
348 problem. This operator is applied to each offspring in the population with a predetermined
349 probability [47]. In our study we used common mutation techniques in a GA. A brief
350 explanation of these techniques is given below.

351

5.6.2.1 Twors mutation

353

354 In Twors mutation [51], the two genes randomly chosen to exchange of their position, as
355 shown in figure 7.

356



357

358 **Fig. 7. Twors mutation**

359

360

5.6.2.2 Reverse Sequence Mutation (RSM)

362

363 In the reverse sequence mutation operator, we take a sequence S limited by two positions i
364 and j randomly chosen, such that $i < j$. The gene order in this sequence will be reversed
365 by the same way as what has been covered in the previous operation. Figure 8 shows the
366 implementation of this mutation operator.

367



368

369 **Fig. 8. Reverse Sequence Mutation**

370
371
372
373
374
375
376
377
378
379
380
381
382

5.6.2.3 One point mutation

In one point mutation, data at a particular point is mutated (a particular gene was randomly selected and then it was replaced with a random state from the available set) [48].

5.6.2.4 Centre inverse mutation (CIM)

The chromosome is divided into two sections. All genes in each section are copied and then inversely placed in the same section of a child [51]. Figure 9 shows the implementation of this mutation operator.

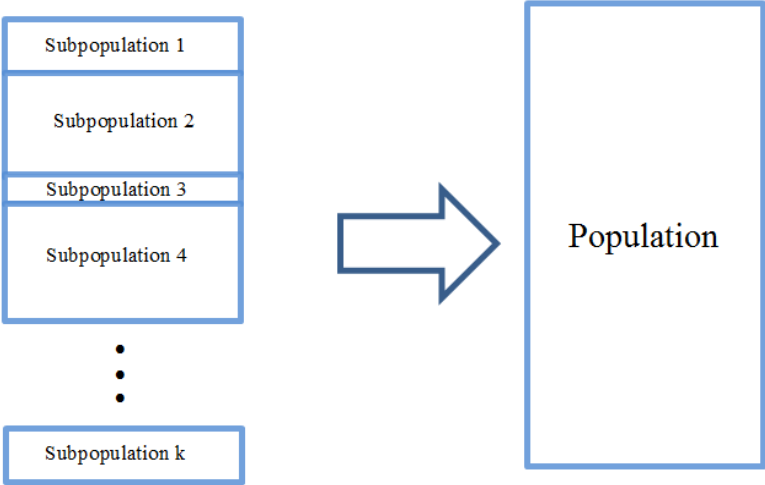


383
384
385
386
387
388
389
390
391

Fig. 9. Centre inverse mutation

5.7 Combination stage

In combination stage, all subpopulations are combined together again to create a new population, as illustrated in figure 10.



392
393
394
395
396
397
398
399
400
401
402
403
404

Fig. 10. Commination stage

5.8 Update the archive of non-dominated solution.

The algorithm maintains an external archive of non-dominated solutions which gets iteratively updated in the presence of new solutions based on the concept of non-domination. The major function of the archive is to store a historical record of the non-dominated solutions found along the heuristic search process [52]. During the implementation of the algorithm a tentative set of non-dominated solutions is stored in archive which gets iteratively updated every generation. Algorithm 1 show the procedure

(Iterative search procedure) which is used to update the archive of non-dominated solutions [53]. The purpose of the function $p^{(t)}$ is to generate a new search point in each iteration t . The function $A^{(t)}$ updates the old archive set $A^{(t-1)}$ using the contents of $p^{(t)}$ and the old archive set $A^{(t-1)}$ and determines the update archive set.

Algorithm 1: Iterative search procedure

```

1.  $t \leftarrow 0$ 
2.  $A^{(0)} = \emptyset$ 
3. While terminate (  $A^{(t)}$  ,  $t$  ) = false do
4.    $t \leftarrow t + 1$ 
5.    $p^{(t)} \leftarrow \text{generate}()$            {generate new search point}
6.    $A^{(t)} \leftarrow \text{update}(A^{(t-1)}, p^{(t)})$  {update archive }
7. End while
8. Output :  $A^{(t)}$ 

```

410
411
412
413
414
415
416
417
418
419
420

6. EXPERIMENTAL RESULTS

In this section, the proposed approach has been applied to multi-objective resource allocation test problems taken from the literature [14, 38] to verify the performance of the proposed algorithm and illustrate its ability for solving such problems. The test problems have been executed on an Intel core i5, 2.6 GHz processor. The proposed approach is coded using MATLAB programming language. The parameter setting used for all runs are depicted in Table 1.

Table 1: The proposed algorithm parameters

Parameter	Values	
	Prolem1	Prolem2
Population size	100	500
Crossover rate	0.65	0.95
Mutation rate	0.02	0.02
Iteration	30	200
Number of cluster (k)	1&4	1&4

421
422
423
424
425
426
427
428

6.1 Test problem 1

Test problem 1 (taken from [14]) of allocating 6 workers to a certain set of 4 jobs. Table 2 provides the expected cost and efficiency.

Table 2: The expected cost and efficiency of test problem 1

Number of worker	JOB							
	1		2		3		4	
	Cost	Efficiency	Cost	Efficiency	Cost	Efficiency	Cost	Efficiency
0	70	0	90	0	85	0	130	0
1	60	25	60	20	60	33	115	13
2	50	42	50	38	50	43	100	24
3	40	55	40	54	55	47	100	32
4	40	63	30	65	40	50	90	39

5	45	69	20	73	30	52	80	45
6	50	74	25	80	25	53	80	50

Tables 3 and 4 present the efficient solution obtained by the proposed algorithm at $k=4$ and $k=1$, respectively. While table 5 presents the results obtained by the effective genetic algorithm [14]. Furthermore figure 11 shows the efficient solution by the proposed algorithm at $k=4$, $k=1$ and the effective genetic algorithm [14].

Table 3: Efficient solution of the proposed algorithm at $k=4$.

Efficient solution				Overall Cost	Overall efficiency
X1	X2	X3	X4		
1	2	1	2	270	120
2	2	1	1	275	126
2	3	1	0	280	129

Table 4: Efficient solution of the proposed algorithm at $k=1$

Efficient solution				Overall Cost	Overall efficiency
X1	X2	X3	X4		
1	2	1	2	270	120
2	2	1	1	275	126
2	3	1	0	280	129

Table 5: Efficient solution of the effective genetic algorithm approach [14]

Efficient solution				Overall Cost	Overall efficiency
X1	X2	X3	X4		
1	2	1	2	270	120
2	2	1	1	275	126
2	3	1	0	280	129

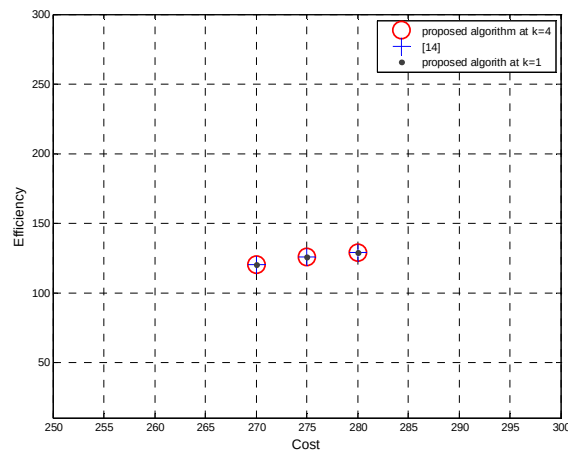


Fig. 11. Efficient solution of test problem 1 obtained by the proposed algorithm at $k=4$, $k=1$ and the effective genetic algorithm approach [14].

It is clear from previous comparison in test problem 1 that the results obtained by the proposed algorithm at k=1, k=4 and the results obtained by effective genetic algorithm is the same values.

6.2 Test problem 2

Test problem 2 (taken from [38]) of allocating 10 workers to a certain set of 4 jobs is solved. Table 6 provides the expected cost and efficiency. Table 7 and table 8 presents the efficient solution obtained by the proposed algorithm at k=4 and k=1, respectively. While Table 9 present the results obtained by multistage decision-based genetic algorithm [38]. Furthermore figure 12 shows the simulation results of the proposed algorithm at different values of k and multistage decision-based genetic algorithm.

Table 6: The expected cost and efficiency of test problem 2

Number of worker	JOB 1		2		3		4	
	Cost	Efficiency	Cost	Efficiency	Cost	Efficiency	Cost	Efficiency
0	41	0	45	0	36	0	46	0
1	38	37	54	49	43	45	78	60
2	46	42	36	55	68	49	88	67
3	32	50	55	59	56	57	64	72
4	78	54	87	62	72	64	90	79
5	76	56	82	67	59	77	80	83
6	72	58	90	73	32	88	120	88
7	84	65	132	80	67	92	104	97
8	80	72	97	87	86	100	96	102
9	92	80	21	95	188	105	86	110
10	96	95	134	102	100	110	120	120

Table 7: Efficient solution of the proposed algorithm at k=4

Efficient solution				Overall Cost	Overall Efficiency
X1	X2	X3	X4		
2	1	6	1	210	239
3	1	6	0	164	187
1	1	6	1	202	234
3	2	2	3	200	226
3	2	1	0	157	150
3	2	1	3	175	222
2	2	6	0	160	185
0	2	6	0	155	143

Table 8: Efficient solution of the proposed algorithm at k=1

Efficient solution				Overall Cost	Overall Efficiency
X1	X2	X3	X4		
2	3	1	0	190	146

3	1	1	0	175	144
3	3	0	1	201	169
3	3	0	0	169	109
3	1	1	1	207	204
3	0	0	0	159	50
3	3	1	1	208	214
3	1	0	0	168	99
3	1	0	1	200	159
3	0	1	1	198	155

Table 9: Efficient solution of multistage decision-based genetic algorithm [38]

Efficient solution				Overall Cost	Overall efficiency
X1	X2	X3	X4		
3	2	1	4	201	229
0	2	6	2	197	210
3	2	5	0	173	182
3	1	6	0	164	175
1	1	6	2	212	241
1	1	5	3	215	235
0	1	6	3	191	209

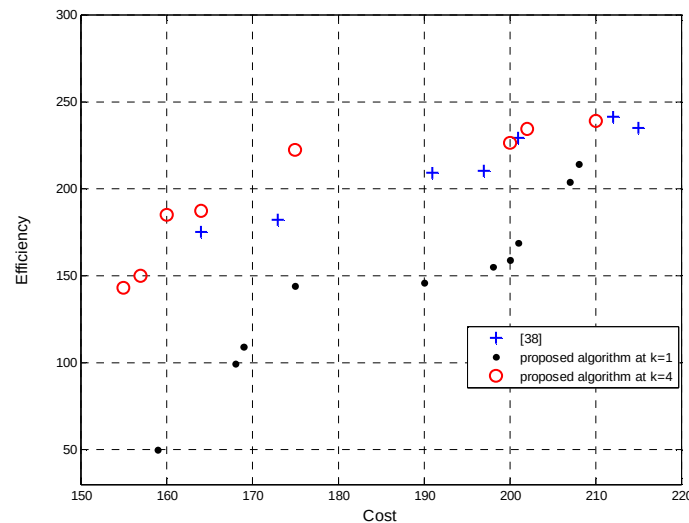


Fig.12. Simulation results of test problem 2 obtained by the proposed algorithm at k=4, k=1 and multistage decision-based genetic algorithm [38].

As shown in figure 12 the results obtained by the proposed algorithm with k=4 is dominated to the most of the results obtained by multistage decision-based genetic algorithm. In addition, the result obtained by the proposed algorithm with clustering technique (i.e. k=4) is dominant to the results obtained by the algorithm without clustering data (i.e. k=1). So the proposed algorithm is more dominant position to be used as an explanation to the problems in hand and different MORAPs.

7. Conclusion

In this paper, GA based on k-means clustering technique is proposed to solve MORAPs. K-means clustering technique divided population of GA into a specific number of subpopulations with dynamic size; where different GA operators can be applied to subpopulations instead of one GA operators applied to the whole population. Furthermore, using GA with K-means clustering preserve and introduce diversity, and enable the algorithm to avoid local minima by preventing the population of chromosomes from becoming too similar to each other and to benefit from the advantages of both types of algorithms. Two test problems taken from the literature are solved to verify the performance of the proposed algorithm and its capability to solve MORAP. The simulation results prove superiority of the proposed algorithm to those reported in the literature, where it completely better than the other approaches.

References

- [1] Ponz-Tienda JL, Yepes V, Pellicer E, Moreno-Flores J. The Resource Leveling Problem with multiple resources using an adaptive genetic algorithm. *Automation in Construction*. 2013;29:161-172.
- [2] Chaharsooghi SK, Kermani AHM. An effective ant colony optimization algorithm(ACO) for multi-objective resource allocation problem (MORAP). *Applied Mathematics and Computation*. 2008; 200:167-177.
- [3]. Lilien GL, Rangaswamy A, Bruggen GHV, Stark K. DSS Effectiveness in Marketing Resource Allocation Decisions: Reality vs. Perception. *Information Systems Research*. 2004;15(3):216-235.
- [4] Wang X, Giannakis GB. Resource Allocation for Wireless Multiuser OFDM Networks. *IEEE*. 2011;57(7):4359-4372.
- [5] Tan S, Weng W, Fujimura S. Scheduling of Worker Allocation in the Manual Labor Environment with Genetic Algorithm. *International Multi Conference of Engineers and Computer Scientists (IMECS) I*; 2009.
- [6] Ehrgott M, Klamroth K, Schwehm C, An MCDM approach to portfolio optimization. *European Journal of Operational Research*. 2004;155(3):752-770.
- [7] Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. *Health Economics*. 2011;20(1):2-15.
- [8] Ibaraki T, Katoh N. *Resource Allocation Problems: Algorithmic Approaches*. MIT Press. Boston(USA); 1988.
- [9] Zhao ZJ, Lau HC, Ge SS. Integrated Resource Allocation and Scheduling in a Bidirectional Flowshop with Multimachine and COS Constraints, *Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*. 2009;39(2):190-200.
- [10] Demeulemeester E, Herroelen W. A branch-and-bound procedure for the multiple resource constrained project scheduling problem. *Management Science*. 1992;38(12):1803-1818.
- [11] Gavish B, Pirkul H. Algorithms for multi-resource generalized assignment problem. *Management Science*. 1991;37(6):695-713.
- [12] Moselhi A, Lorterapong P. Least impact algorithm for resource allocation. *Canadian Journal of Civil Engineering*. 1993;20(2):180-188.
- [13] Deb K. *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Son. Chichester (UK); 2001.

- 541 [14] Osman MS, Abo-Sinna MA, Mousa AA . An effective genetic algorithm approach to
542 multi-objective resource allocation problems. *Applied Mathematics and Computation*.
543 2005;163:755–68.
- 544 [15] Wang KJ, Lin YS. Resource allocation by genetic algorithm with fuzzy inference. *Expert*
545 *Systems with Applications*.2007;33:1025–1035.
- 546 [16] Sousa T, Vale Z, Carvalho JP, Pinto T, Morais H. A hybrid simulated annealing
547 approach to handle energy resource management considering an intensive use of
548 electric vehicles. *Energy*. 2014;67:81-96.
- 549 [17] Liang YC, Chuang CY. Variable neighborhood search for multi-objective resource
550 allocation problems. *Robotics and Computer-Integrated Manufacturing*. 2013; 29: 73–
551 78.
- 552 [18] Wu J, Chang Z, Yuan L, Hou Y, Gong M. A Memetic Algorithm for Resource Allocation
553 Problem Based on Node-Weighted Graphs. *IEEE Computational Intelligence*
554 *magazine*. 2014; 9(2): 58–69.
- 555 [19] Jia Q, Seo Y. An improved particle swarm optimization for the resource-constrained
556 project scheduling problem. *International Journal of Advanced Manufacturing*
557 *Technology*. 2013; 67: 2627–2638.
- 558 [20] Tounsi M. A Heuristic-Based Technique for University Resource Allocation Problems.
559 2006 IEEE GCC conference. 2006;1-6.
- 560 [21] Qiao F, Lin P. A Differential Evolutionary Based Algorithm for Multiuser OFDMA
561 System Adaptive Resource Allocation. 4th International Conference on Wireless
562 Communications, Networking and Mobile Computing (WiCOM '08). 2008;1-5.
- 563 [22] Yang B , Hu Y, Huang C-Y. An Architecture-Based Multi-Objective Optimization
564 Approach to Testing Resource Allocation. *IEEE Transactions on Reliability*.
565 2014;PP(99):1-19.
- 566 [23] Fatima S , Bader-El-Den M.. Co-evolutionary Hyper-Heuristic Method for Auction Based
567 Scheduling. *IEEE Congress on Evolutionary Computation (CEC2010)*. 2010;1-8.
- 568 [24] Zhao SY, Szeto KY. Combining Exhaustive Search with Evolutionary Computation via
569 Computational Resource Allocation. *IEEE Congress on Evolutionary Computation*
570 *(CEC 2007)*. 2007;1878-1881
- 571 [25] Murakami K, Gen M, Tasan SO, Oyabu T. A Solution of Human Resource Allocation
572 Problem in a Case of Hotel Management. 40th International Conference on Computer
573 and Industrial Engineering (CIE). 2010;1-6.
- 574 [26] Yan HC, Liu HL, Wang Q. Dynamic Resource Allocation of TD-LTE System Based on
575 Improved Quantum Evolutionary Algorithm. 9th International Conference on
576 Computational Intelligence and Security (CIS). 2013;26-30.
- 577 [27] Verma M, Srivastava M, Chack N, Diswar AK, Gupta N. A Comparative Study of
578 Various Clustering Algorithms in Data Mining. *International Journal of Engineering*
579 *Reserch and Applications (IJERA)*. 2012; 2(3):1379-1384.
- 580 [28] Aggarwal CC, Reddy CK. *DATA CLUSTERING Algorithms and Applications* . CRC
581 Press Taylor & Francis Group. New York(USA); 2014.
- 582 [29] MacQueen JB. Some Methods for classification and Analysis of Multivariate
583 Observations. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics*
584 *and Probability*, Berkeley, University of California Press. 1967:281-297.
- 585 [30] Arai K, Bu XQ. ISODATA clustering with parameter (threshold for merge and split)
586 estimation based on GA: Genetic Algorithm. *Reports of the Faculty of Science and*
587 *Engineering, Saga University*. 2007;36(1):17-23.
- 588 [31] Ng RT, Han J. CLARANS: A Method for Clustering Objects for Spatial Data Mining.
589 *IEEE Transactions Knowledge Data Engineering*.2002; 14(5): 1003 - 1016
- 590 [32] Judd D, McKinley PK, Jain AK. Large-scale parallel data clustering. *IEEE Transactions*
591 *on Pattern Analysis and Machine Intelligence*, 1998;20(8):871-876.
- 592 [33] Edlaa DR, Janaa PK. A prototype-based modified DBSCAN for gene clustering,
593 *Procedia Technology*. 2012;6:485–492.

- [34] Rani Y, Manju , Rohil H . Comparative Analysis of BIRCH and CURE Hierarchical Clustering Algorithm using WEKA 3.6.9 . The SIJ Transactions on Computer Science Engineering & its Applications (CSEA).2014; 2(1):25-29.
- [35] Abu-Srhan A, Al Daoud E. A Hybrid Algorithm Using a Genetic Algorithm and Cuckoo Search Algorithm to Solve the Traveling Salesman Problem and its Application to Multiple Sequence Alignment International Journal of Advanced Science and Technology. 2013;61:29-38.
- [36] Goldberg DE. Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley Publishing Company .Boston (USA); 1989.
- [37] Michalewicz Z. Genetic Algorithms+Data Structures¼Evolution Programs, 3rd ed. Springer. Verlag ;1996.
- [38] Lin CM, Gen M. Multi-objective resource allocation problem by multistage decision-based hybrid genetic algorithm. Applied Mathematics and Computation. 2007;187:574-583.
- [39] Mohd RN, Geraghty J. Genetic Algorithm Performance with Different Selection Strategies in Solving TSP. Proceedings of the World Congress on Engineering II; 2011.
- [40] Sakthi M, Thanamani AS. An Effective Determination of Initial Centroids in K-Means Clustering Using Kernel PCA . International Journal of Computer Science and Information Technologies. 2011;2(3):955-959.
- [41] Jain A. Data clustering: 50 years beyond k-means. Pattern Recognition Letters. 2010; 31(8):651–666.
- [42] Miettinen K. Nonlinear Multi-objective Optimization. Kluwer Academic Publishers. Boston(USA); 1999.
- [43] Steuer RE. Multiple Criteria Optimization: Theory, Computation, and Application. John Wiley & Sons. New York(USA); 1986.
- [44] Srinivas N, Deb K. Multi-objective optimization using non-dominated sorting in genetic algorithms. Evolutionary Computation. 1999;2(3):221–248.
- [45] Murata T, Ishibuchi H, Tanaka H. Multiobjective genetic algorithm and its application to flowshop Scheduling. Computers and Industrial Engineering. 1996;30(4):957–968.
- [46] Sivaraj R, Ravichandran T. a review of selection methods In genetic algorithm. International Journal of Engineering Science and Technology. 2011;3(5): 3792-3797.
- [47] Tzeng GH, Huang JJ. Fuzzy Multiple Objective Decision Making. Chapman and Hall/CRC Press; 2013.
- [48] Kaya M. The effects of two new crossover operators on genetic algorithm performance. Applied Soft Computing. 2011;11:881–890.
- [49] Gen M, Cheng R. Genetic Algorithms and Engineering Optimization. John Wiley & Sons. New York(USA); 2000.
- [50] Faraji R, Naji HR. An efficient crossover architecture for hardware parallel implementation of genetic algorithm. Neurocomputing. 2014;128:316–327.
- [51] Abdoun O, Abouchabaka J, Tajani C. Analyzing the Performance of Mutation Operators to Solve the Travelling Salesman Problem. international Journal of Emerging Sciences. 2012;2(1):61-77.
- [52] Goh CK, Ong YS, Tan KC. Multi-Objective Memetic Algorithms. springer, Berlin (Germany); 2009.
- [53] Laumanns M, Thiele L, Deb K, Zitzler E. Archiving with Guaranteed Convergence And Diversity in Multi-objective Optimization. In GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference, Morgan Kaufmann Publishers, New York (USA); 2002:439-447.