Computing the Most Significant Solution from Pareto Front obtained in Multi-objective Evolutionary

Mrs.P.M.Chaudhari
PhD Student ,
Post Graduate Department of
Computer Science & Engineering,
G. H. Raisoni College of Engineering,
Nagpur,India

Dr. R.V. Dharaskar
Professor & Head,
Post Graduate Department of
Computer Science & Engineering,
G. H. Raisoni College of
Engineering, Nagpur,India

Dr. V. M. Thakare
Professor & Head,
Post Graduate Department of
Computer Science,Faculty of
Engineering &Technology,
S.G.B. Amravati University,
Amravati.

Abstract- **Problems with multiple objectives can be solved by using Pareto optimization techniques in evolutionary multi-objective optimization algorithms. Many applications involve multiple objective functions and the Pareto front may contain a very large number of points. Selecting a solution from such a large set is potentially intractable for a decision maker. Previous approaches to this problem aimed to find a representative subset of the solution set. Clustering techniques can be used to organize and classify the solutions. Implementation of this methodology for various applications and in a decision support system is also discussed.**

Keywords- Multiobjective,Pareto front ,Clustering techniques

I. INTRODUCTION

Multi-objective optimization is applied to a variety of fields and sufficient computational power exists to generate very large non-dominated sets for these problems. In order to be sufficiently representative of the possibilities and tradeoffs, a non-dominated set may be too large for decision makers to reasonably consider. Some means of reducing or organizing the non-dominated set is needed [1]. Several researchers have dealt with this issue using cluster analysis or filtering. This paper differs from their work in that it intends to not only make the non-dominated set tractable but to do so without removing any elements of the non-dominated set before presenting the solutions to the decision makers.

Cluster analysis can be applied to the results of a multi-objective optimization algorithm to organize or partition solutions based on their objective function values. The goal of clustering is to create an "efficient representation that characterizes the population being sampled" [2]. Such a representation allows a decision maker to further understand the decision by making available the attainable limits for each objective, key decisions and their consequences, and the most relevant variables; this presentation is an improvement on a list of potential solutions and their associated objective function values.

This paper details a k-means cluster analysis approach.

II. MULTIOBJECTIVE OPTIMIZATION AND BACKGROUND DETAILS

Three approaches can be taken to find a solution to multi-objective problems (Benson and Sayin 1997). The first approach entails reformulating the problem as a single objective problem. To do so additional information is required from the decision makers such as the relative importance or weights of the objectives, goal levels for the objectives, values functions, etc. The second approach requires that the decision makers interact with the optimization procedure typically by specifying preferences between pairs of presented solutions. The third approach, Pareto optimization, finds a representative set of non-dominated solutions approximating the Pareto front. Pareto optimization methods, such as evolutionary multi-objective optimization algorithms, allow decision makers to check the potential solutions without a priori judgments regarding the relative importance of objective functions. Post-Pareto analysis is necessary to select a single solution for implementation.

All three approaches to solving multi-objective optimization problems have shortcomings. The solution returned by the single objective approach can be highly dependent on the weights and, in non-convex problems, the responses to changes in weights or goals may be unpredictable. As well, with conflicting and non-commensurate criteria it can be hard to make value judgments such as choosing weights or goals for the criteria. Given decision maker input the first approach returns a single solution. Interactive approaches consider only a small set of non-dominated solutions due to the effort required [3]. Pareto optimization approaches return a potentially large number of solutions for consideration. Selecting a single solution from a large non-dominated set is likely to be difficult for any decision maker. It was proposed that an ideal solution procedure for multi-objective optimization is to provide the decision makers with a globally representative subset of the non-dominated set that is sufficiently small so as to be tractable [4]. This work aims to approach this ideal procedure by accepting the computational effort required for generating a large non-dominated set and subsequently organizing it based on its structure. This approach allows decision makers to tractably consider interesting subsets without a priori removal of any solutions from consideration.

Any Pareto optimization method could be employed in this methodology. Evolutionary multiobjective algorithms apply biologically inspired evolutionary processes as heuristics to generate non-dominated sets of solutions. It should be noted that the solutions returned by evolutionary multiobjective algorithms may not be Pareto optimal, that is, globally non-dominated, but the algorithms are designed to evolve solutions that approach the Pareto front and spread out to capture the diversity existing on the Pareto front in order to obtain a good approximation of the Pareto front.

III. POST-PARETO ANALYSIS

Post-Pareto analysis aids decision makers in choosing a single solution from the potentially large set of Pareto optimization results. Several researchers have applied clustering methods in different ways to non-dominated sets to aid decision makers. Most of these methods use the similarity of elements in the non-dominated set based on their objective function values and remove elements that are too similar to other elements.

The main goal of multi-objective optimization is to seek Pareto-optimal solutions. Over the years there have been various approaches toward fulfillment of this goal. It has been observed that convergence and diversity are two conflicting criteria which must be balanced in trying to generate the entire efficient front [5]. Clearly, there are two different possible principles for generating a set of solutions representing the entire Pareto-optimal front:

- One-at-a-time strategy, and

- Simultaneous strategy

In the former method, a multi-objective optimizer may be applied one at a time with the goal of finding one single Pareto-optimal solution. Most classical generating multi-objective optimization methods use such an iterative scalarization scheme of standard procedures. The main criticism of most of these approaches is that although there are results for convergence, diversity among obtained Pareto-optimal solutions is hard to maintain in the objective space. Moreover, a careful thought suggests that a systematic variation of weight vectors or "parameters in these scalarization techniques does not guarantee a good diversity in the solution sets [6]. Another important issue is that independent applications of a single-objective optimization algorithm to find different Pareto-optimal solutions one-at-a-time do not make an efficient search and the search effort required to solve the problem to optimality this way needs to be found in every single time the algorithm is applied.

Morse (1980) detailed one of the first applications of cluster analysis to a non-dominated set. The multi-objective programs considered were linear programs. A solution was removed from the non-dominated set if it was indistinguishable from another solution based on decision maker-defined thresholds. Morse (1980) evaluated seven hierarchical clustering methods. Ward's method, the group average method, and the centroid method performed very well; the other hierarchical clustering methods considered exhibited chaining which reduced the usefulness of the cluster structure. Ward's method was preferred since the clusters at the same level of the hierarchy were of similar size and shape although it performed only slightly better than the centroid and group average methods (Rosenman and Gero 1985).

Rosenman and Gero (1985) applied complete linkage hierarchical clustering to 'reduce the size of the Pareto optimal set whilst retaining its shape'. This method allowed control of the diameter of the resulting clusters. They distinguished that solutions whose vectors of objective function values are similar may have decision variable vectors that are similar or very different but this idea was not further explored. The objective functions were considered successively in order to avoid the implicit aggregation in applying proximity measures. First, elements of the non-dominated set were clustered using a single criterion. If a solution within a cluster dominated another solution in the cluster on all criteria except the clustering criterion then the dominated solution was eliminated from consideration. The process was repeated for each criterion until the non-dominated set was sufficiently small.

This paper differs from the above wherein partitional (k-means) clustering is used for combinatorial multi-objective problems. Either the most interesting cluster, i.e., the 'knee' cluster, was considered in detail by discarding the solutions in other clusters, or one solution from each of the k clusters was considered to form a representative subset of the non-dominated set.

IV. CLUSTER ANALYSIS

Cluster analysis, also known as unsupervised learning, is one of the most useful methods in the cluster analysis process for discovering groups. Clustering aims to organize a collection of data items into clusters, such that objects within the same cluster have a high degree of similarity, while objects belonging to different clusters have a high degree of dissimilarity. Cluster analysis makes it possible to look at properties of whole clusters instead of individual objects. This is a simplification that is useful when handling large amounts of data [9].

According to the method adopted to define clusters, the algorithms can be broadly classified into the following types: Partitional and Hierarchical [10]. Partitional clustering attempts to directly decompose the data set into a set of disjoint clusters. Probably, one of the most popular partitional methods is the *k-means clustering* algorithm. The *k-means clustering* algorithm is well known for its efficiency in clustering data sets [11]. The grouping is done by calculating the centroid for each cluster, and assigning each observation to the group with the closest centroid. For the membership function, each data point belongs to its nearest center, forming a partition of the data. A recurrent problem that many clustering algorithms encounter is the choice of the number of clusters. Thus, different cluster validity indices have been suggested to address this problem, since this is an important issue for partitional clustering in general. A cluster validity index indicates the quality of a resulting clustering process.

The silhouette plot method is one of these cluster validity techniques [12]. Then, the clustering partition that optimizes the validity index under consideration is chosen as the best partition. The silhouette plot is used to evaluate the quality of a clustering allocation, independently of the clustering technique that is used [13].

V. METHODOLOGY

Proposed Approach

This approach is proper for decision-makers that do not have a prior knowledge of the relative importance of the conflicting objectives in multi-objective optimization problem.

The developed approach is based on the following steps:

1. Obtain the entire Pareto-optimal set or sub-set of solutions by using a multiple-objective evolutionary algorithm (MOEA) or by another means.

2. Apply the *cluster analysis* algorithm to form clusters on the solutions enclosed in the Pareto set.

3. To determine the "optimal" number of clusters, k, in this set, silhouette plots are used. A value of the silhouette width, $s(i)$, is obtained for several values of k. The clustering with the highest average silhouette width is selected as the "optimal" number of clusters in the Pareto-optimal set.

4. For each cluster, select a representative solution. To do this, the solution that is closest to its respective cluster centroid is chosen as a good representative solution.

5. Analyze the results. At this point, the decision-maker can either:

5.1 Analyze the "knee" cluster. The suggestion is to focus on the cluster that has solutions that conform to the "knee" region. The "knee" is formed by those solutions of the Pareto-optimal front where a small improvement in one objective would lead to a large deterioration in at least one other objective. Moreover, from this "knee" cluster the decision maker can select a promising solution for system implementation. This would be the solution nearby to the ideal or utopian solution of the multiple objective problems, in a standardized space.

5.2 Analyze the k representative solutions and/or select the most promising solutions among this k set, selecting the solution closest to the ideal point. By applying the proposed technique, the Pareto-optimal front of a multiple objective problem can be reduced to the "knee cluster" as in 5.1, or to a set of k solutions as in 5.2. In both cases the decision maker can choose a superior tradeoff for system implementation by selecting the closest solution to the ideal or utopian solution of the multiple objective problems, in a standardized space.

A Matlab code is developed to perform the steps of the proposed method. From standardized data, the code will run the *clustering* algorithm and from two to a specified number of means it will calculate the average silhouette values and it will return the value of k suggesting the most optimal allocation. After this, it will also return the "knee cluster" of the optimal partition, the k representative solutions of the Pareto front, and in both cases, the solution closest to the ideal or utopian point.

Multi-Objective Redundancy Allocation Problem

A R.A.P. example was solved to exemplify how data clustering can be of great aid for the decision-maker. The example system configuration consists of 3 subsystems, with an option of 5, 4 and 5 types of available components for each subsystem, respectively. The optimization involves selection from among these component types. The maximum number of components is 8 for each subsystem. Table 1 defines the component choices for each subsystem.

TABLE 1. Component selection for each subsystem.

Design Alternative j	Subsystem i								
	1			2			3		
	R	C	W	R	C	W	R	C	W
1	0.94	9	9	0.97	12	5	0.96	10	6
2	0.91	6	6	0.86	3	7	0.89	6	8
3	0.89	6	4	0.70	2	3	0.72	4	2
4	0.75	3	7	0.66	2	4	0.71	3	4
5	0.72	2	8				0.67	2	4

Another way to take advantage of this method is to consider the cluster(s) that contain(s) the most interesting solutions of the Pareto-optimal set, i.e., those where a small improvement in one objective would lead to a large deterioration in at least one other objective. These solutions are often referred as "knees." In this case, as we can see from Figure 4, solutions in cluster 4 are likely to be more significant to the decision-maker. The maximum and minimum values of reliability, cost and weight of cluster 4 are shown in Table 2.

TABLE 2. Maximum and minimum values in cluster 4.

Cluster 4	Reliability	Cost	Weight
max	0.999036	77	55
min	0.961883	38	32

At this point, the decision-maker has two choices: either to choose solution #43 from cluster 4 as a good representative solution of this "knee" region or decide to focus his/her search more intensely on this knee region. Then, the initial 75 solutions have been reduced to only the 30 solutions found in the "knee" region as shown in Figure 1.
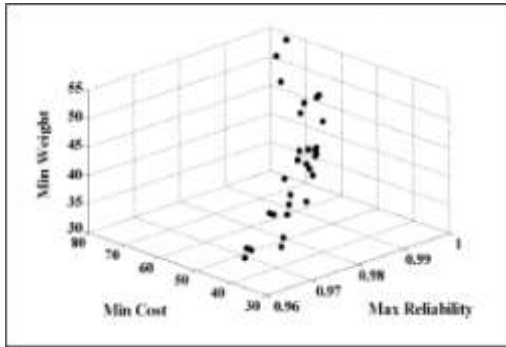
Figure 1. "Knee" of the Pareto-optimal set.

For instance, if the decision-maker decides to further investigate this "knee" region, then the 30 solutions contained in cluster 4 are further investigated. Clustering is again used to find groups just on this reduced space, and with the use of the silhouette plots, 11 was found to be the optimal number of clusters. In this way, one systematically contracts the subspace in the direction of the most relevant solutions for the decision-maker. Figure 2 shows the clusters found on the original cluster 4 from normalized data.
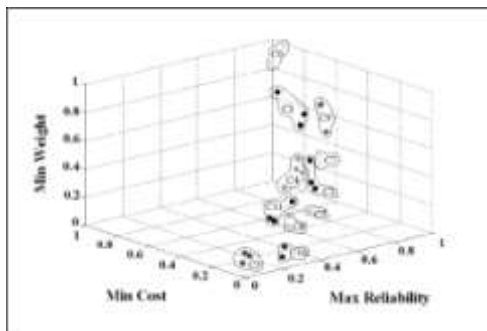

Figure 2. Clusters from original cluster 5.

Since the original cluster 5 already contained promising trade-offs, plotting the solutions in two dimensions can be of graphical support for the decision-maker. Figures 3, 4 and 5 plot reliability vs. cost, reliability vs. weight and cost vs. weight, respectively, from normalized (0 to 1) objective function data.
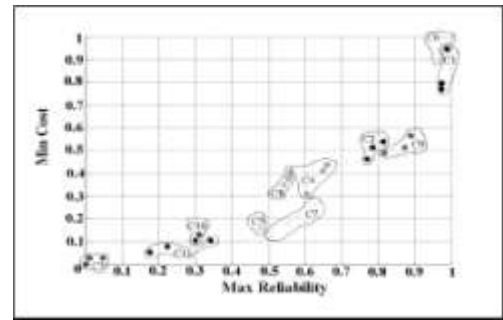

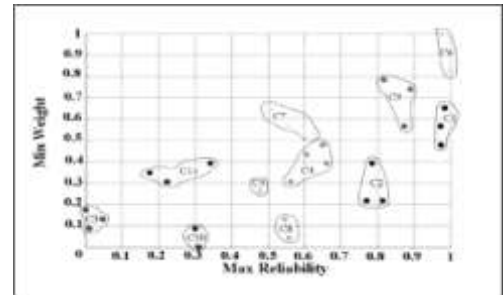Figure 3.   Reliability vs. Cost.

Figure 4. Reliability vs. Weight.

From Figures 3 and 4, clusters 3, 10 and 11, can be considered as undesirable because they do not have a large reliability compared with the other clusters, but in Figure 5, it can be observed that these three clusters are the ones that provides the minimum values for cost and weight. Nevertheless, clusters 1 and 6 in Figures 3 and 4 have large reliability but it is achieved at comparatively high cost and weight.

The analysis of these trade-offs continues until a solution or a small portion of the nondominated set is located. Then, this solution or sub-set will contain the preferred solutions of the overall problem. It is important to note that, even when the space has been reduced to the "knee" region or to the region that contains the most promising solutions, in the absence of information, none of the corresponding trade-offs can be said to be better than the others. Thus, the choice of one solution over the other is going to lie on the capability of the decision-maker and on his/her knowledge of the system's intended usage and the priorities and preferences of the system's intended user/owner.
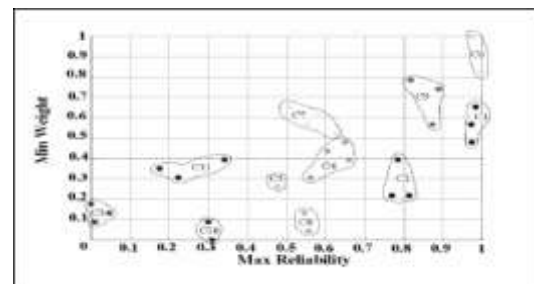

Figure 5. Cost vs. Weight

The representative solutions of these 11 clusters found on the "knee" region are shown in Table 3 with their corresponding values of reliability, cost and weight.

TABLE 3. Clustering results of the "knee"

	# of solutions	Representative solution	Reliability	Cost	Weight
Cluster 1	3	51	0.998043	68	45
Cluster 2	3	47	0.992115	59	37
Cluster 3	3	28	0.963644	39	35
Cluster 4	4	44	0.986416	55	41
Cluster 5	1	34	0.979653	44	38
Cluster 6	2	56	0.999036	77	51
Cluster 7	4	40	0.983483	46	46
Cluster 8	2	36	0.982178	52	35
Cluster 9	3	50	0.994940	60	49
Cluster 10	2	31	0.973035	42	34
Cluster 11	3	30	0.970198	41	39

For this particular multi-objective RAP, clusters 2, 4 and 7 seem to contain desirable solutions. For ease of interpretation and analysis, the 11 representative solutions of this "knee" region are plotted in two dimensions in Figure 6 for reliability vs. cost. From this figure, one can easily notice that solutions #47, #44 and #40 belonging to clusters 2, 4 and 7 respectively are the ones that are presented to the decision-maker as good trade-offs if no previously defined objective function preference have been specified by the decision-maker.

For example, solution #44, shown in Figure 7, achieves a reliability of 0.986416 at a cost of 55 and a weight of 41. For system execution, the configuration is composed of one component of type 1, one component of type 2 and one component of type 5 for subsystem 1; two components of type 1 for subsystem 2, and one component of type 1 and one component of type 3 for subsystem 3.

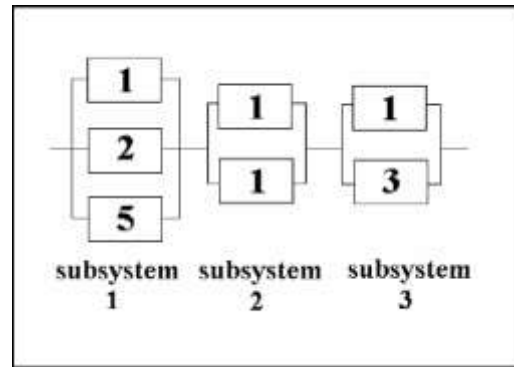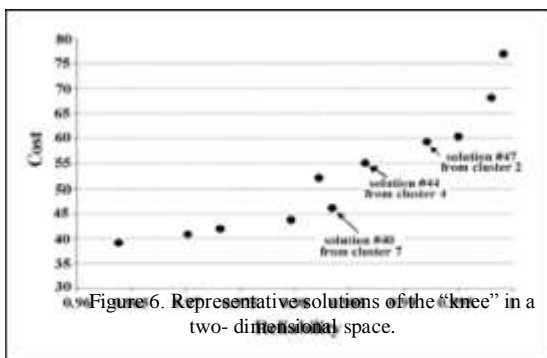Figure 6. Representative solutions of the "knee" in a two-dimensional space.

Figure7. The system configuration for solution #44.

VI. CONCLUSION

Pareto optimization methods allow the use of multi-objective optimization models without prior decision maker preferences. The decision makers can consider the possibilities and trade-offs between objectives before selecting a solution for implementation. These methods undergo the limitations of requiring the decision makers to consider many possible solutions resulting from the optimization procedure. This paper developed and evaluated a cluster analysis methodology to address the issue.

REFERENCES

[1] Balling, R. 2004. Applications of Multi-Objective Evolutionary Algorithms: Advances in Natural Computation, Volume 1, City and Regional Planning via a MOEA: Lessons Learned. World Scientific, Singapore, 227–245.

[2] Benson, H.P. and S. Sayin. 1997. Towards finding global representations of the efficient set in multiple objective mathematical programming. Naval Research Logistics, 44, 47–67.

[3] Cormack, R. M. 1971. A review of classification. Journal of the Royal Statistical Society A, 134(3), 321– 367.

[4] Deb, K., A. Pratap, S. Agarak, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197.

[5] Dubes, R. and A.K. Jain. 1976. Clustering techniques: The user's dilemma. Pattern Recognition, 8, 247– 260.

[6] Dubes, R. and A.K. Jain. 1979. Validity studies in clustering methodologies. Pattern Recognition, 11, 235–254.

[7] Dubes, R.C. 1993. Handbook of Pattern Recognition and Computer Vision, Cluster Analysis and Related Issues. World Scientific Publishing Company, Salem, Massachusetts, 3–32.

[8] Gnanadesikan, R. 1995. Weighting and selection of variables for cluster analysis. Journal of Classification, 12, 113–136.

[9] Jain, A. K., M. N. Murty, and P. J. Flynn. 1999. Data clustering: A review. ACM Computing Surveys, 31(3).

[10] Jain, Anil K. and Richard C. Dubes. 1988. Algorithms for clustering data. Prentice Hall, Englewood Cliffs, NJ.

[11] Mattson, C.A., A.A. Mullur, and A. Messac. 2004. Smart pareto filter: Obtaining a minimal representation of multiobjective design space. Engineering Optimization, 36(6), 721–740.

[12] Milligan, G.W. and M.C. Cooper. 1988. A study of standardization of variables in cluster analysis. Journal of Classification, 5, 181–204.

[13] Morse, J. N. 1980. Reducing the size of the nondominated set: Pruning by clustering. Computers and Operations Research, 7, 55–66.

[14] Roberts, S. A. 2003. Configuration optimization in socio-ecological systems. Ph.D. thesis, Department of Systems Design Engineering, University of Waterloo, Waterloo, ON.

[15] Rosenman, M. A. and J. S. Gero. 1985. Reducing the pareto optimal set in multicriteria optimization (with applications to pareto optimal dynamic programming). Engineering Optimization, 8, 189– 206.

[16] Seo, J. and B. Shneiderman. 2002. Interactively exploring hierarchical clustering results. Computer, 35(7), 80–86.

[17] Small, C.G. 1996. The Statistical Theory of Shape. Springer, New York.

[18] Taboada, H., F. Baheranwala, D. Coit, and N. Wattanapongsakorn. 2007. Practical solutions for multi-objective optimization: An application to system reliability design problems. Reliability Engineering and System Safety, 92(3), 314–322.

[19] Ward, J.H., Jr. 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58(301), 236–244.

[20] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[21] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[22] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[23] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[24] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

AUTHORS PROFILE

Pallavi Chaudhari is currently an Assistant professor of Information Technology at Priyadarshini Institute of Engineering & Technology, Nagpur since 2004. She has obtained her M.E. in Computer Engineering in 2006 from University of Pune, India. She is pursuing Ph.D. in Computer Science. Her research interests are mainly in Genetic Algorithms, Artificial Intelligence, Clustering Techniques and their applications.

Dr. Rajiv Dharaskar is presently working as Professor and Head of PG Department of Computer Science and Engineering at G.H. Raisoni College of Engineering, Nagpur. He is Ph.D. in Computer Science & Engineering in the Faculty of Engineering & Technology, M.Tech. in Computers, P.G. Dip., M.Phil., and M.Sc. He is having 24 years of teaching

and 18 years of R&D experience in the field of Computers & IT. He is approved PhD guide for Computer Engineering and Science for Nagpur and Amravati University and 22 research scholars are perusing Ph.D. degree under his guidance. He is an author of number books on Programming Languages**.**

Dr V M Thakare is Professor and Head of PG department of computer Science and Engg in SGB Amravati University Amravati, Maharastra (India) and has completed ME in Advance Electronics and Ph.D. in computer Science & Engg. His Areas of Research are Robotics and Artificial Intelligence, Information Technology. He is Recognized Guide for Computer Science and Computer Engineering in this University and in other universities also. He has also received national level excellent paper award. More than 10 candidates are working for Ph D Under his supervision. He has Published and presented more than 115 papers at National and International level. He has worked on various national level bodies like AICTE,UGC and also worked on various bodies of other Universities. He is presently a member of BOS, RRC, BUTR of this university and also chairman and member of various committees of this university.