International Journal of Computer and Communication Engineering, Vol. 2, No. 2, March 2013

Analysis of Population Based Metaheuristic Used for Gene Clustering

Arpit Jain, Shikha Agrawal, Jitendra Agrawal, and Sanjeev Sharma

***Abstract*—Gene Clustering is one of the most popular application in the field of Bioinformatics. It is a method of grouping gene into clusters, such that each cluster must have similar gene expression levels. The two most popular population based globalized search algorithms Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) are used for clustering gene expression data but the main drawback of these two algorithm's is that they got trapped in local optima while exploring across the problem space. So the PSO and GA has been hybridized with several other methods, to overcome the local optima problem. This paper presents a literature survey of both GA and PSO's application in Gene Clustering. PSO and GA variants and their hybridization with other metaheuristics are also described in this paper. An attempt is made to provide a guide for the researchers who are working in the area of population based gene clustering.**

***Index Terms*—Genetic Algorithm, K-means, Particle Swarm Optimization, SVM.**

I. INTRODUCTION

Bioinformatics field deals with the analysis of the biological data, which stored information in the form of DNA, protein sequence in various biological databases. The Bioinformatics field involves various problem such as Gene Clustering Problem, Molecular Docking Problem, Multiple Sequence Alignment Problems, Phylogenetic tree construction , RNA Secondary Structure Prediction, Protein Secondary Structure Prediction, Fragment Assembly Problem. In Gene Clustering the gene is defined as a sequence of nucleotide bases, which carry the required information for the purpose of protein formation, which helps in the construction of the structural components. Clustering is a process of combining input patterns into groups based on the similar functionality. So, Gene clustering is a process of grouping genes in the same clusters based on similar gene expression levels. Gene expression refers to a process through which the coded information from a gene is converted into structures operating in the cell. In order to cluster the Gene expression data , various population based algorithms were used by a number of researchers. The two most popular algorithms among these are PSO (Particle Swarm Optimization) and GA (Genetic Algorithm).

II. GENETIC ALGORITHM BASED GENE CLUSTERING

A. Introduction

Genetic Algorithm are heuristic search algorithm based on the technique inspired by natural evolution, such as inheritance, mutation, selection,and crossover. The evolution usually starts from pool of population .Initially, the fitness value of every individual is calculated. Then multiple individuals are selected from the current population based on their fitness value which are recombined to form a new population. The new population is then used in the next iteration of the algorithm. The algorithm usually terminates when a solution that produce an output is close enough or equal to the desired answer or satisfactory fitness level has been reached for the population. However GA find application in Bioinformatics, Phylogenetics, Engineering, Chemistry, Manufacturing, Mathematics, Physics, Computational Science and other fields. From the literature survey it has been found that a genetic algorithm can be applied in following two ways to solve the gene clustering problem:

1) For Gene selection and classification
2) For optimization of gene clustering process.

The next section B and C describes the work done by several researchers in the above mentioned category.

B. For Gene Selection and Classification

In order to classify the Gene expression data Li.et.al [1] proposed an algorithm that combines the Genetic Algorithm and the K-nearest neighbor (i.e. KNN) method. The method is a multivariate approach in which several genes were analyzed. Many subsets of genes were obtained based on the frequency of the occurrence of the genes and thus a predictor set is formed. The purpose of the predictor is to classify the samples. The GA/KNN method maintains the heterogeneity within the classes, which help in the exploration of the subclasses. They applied the algorithm for gene expression data for normal versus tumor tissue for human colon. The two distinct clusters were observed when the 50 most frequently selected genes were used to classify all of the samples in the data sets studied and they observe that the majority of samples were classified correctly based on the proposed algorithm.

When we classify the large scale gene expression data the main problem occurs when the expression dataset which is analyzed involve multiple classes. In such circumstances one needs to use an efficient selection of good predictive gene groups, so that the gene expression data could be classified easily. In order to solve the above problem Ooi. et.

Manuscript received September 10, 2012; revised October 8, 2012.

Arpit Jain, Jitendra Agrawal, and Sanjeev Sharma are with School of Information Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya Bhopal, India (e-mail: arpitcs09@gmail.com, jitendra@rgtu.net, sanjeev@rgtu.net).

Shikha Agrawal is with University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya Bhopal, India (e-mail: shikha@rgtu.net).

al [2] proposed an algorithm that combines the maximum likelihood (MLHD) classification method with Genetic Algorithm. The GA/MLHD classification strategy involves two components. In the first component a GA-based gene selector is used to find a set of R genes which is used to classify the samples, where R lies in the range of [Rmin, Rmax]. The actual classification process is performed using second component where each individual in the population represents a specific gene predictor subset, and to determine the classification accuracy of the prediction set, a fitness function is used. The experimental result shows that GA/MLHD-based approach achieves high classification accuracies than other predictive methods on the same multi-class test datasets.

Sihua *et al.* [3] proposed an algorithm that hybridizes the methodologies of GA, AP (All paired/SVM (support vector machine) [4] and RFE (Recursive feature elimination) [5] for classification of the array-based multiclass tumor for cancer diagnosis. The algorithm involves a pre-filtering process of various cancer types like binary tumor classifications via AP/SVM classifier, a voting scheme to go from binary to multiclass classification based upon AP/SVM results, GA feature selection and multiclass classification optimization via LOOCV fitness test, RFE through AP/SVMs and LOOCV test to further eliminate the non predictive features in the GA-derived gene set. They have tested the algorithm on NCI60 and GCM data set and found that the algorithm yielded accuracies of 87.93% for the eight classes and 85.19%for the fourteen class cancer classifications and thus outperforms previously existing methods.

Although the microarray technology has widely used for the study of gene expression in cancer diagnosis, but it has a drawback that it does not obey some of the assumptions of parametric statistical methods, such as the assumption of normal distribution or independent and microarray data contain thousands of genes, so the type I error will be more expanded. Therefore the Lee. et. al [6] proposed a gene selection method without assumption restriction for reducing the dimension of the data set. They apply an adaptive genetic algorithm/ k-nearest neighbor (AGA/KNN). The reason for using such a combination of AGA and KNN is that AGA is a search tool of machine learning and imitates the biological system to find the near optimal solution, so it is suitable to analyze high-dimensional, noisy data. Second, KNN is one of the most widely used classification technique due to its simplicity and effectiveness. Each sample can be classified according to the classification of its k nearest neighbors which are determined by Euclidean distance. Author found that AGA/KNN reduces the dimension of the data set, and all the test samples were classified correctly. The accuracy of AGA/KNN is higher than that of GA/KNN, and it only takes half of the CPU time as compared to GA/KNN.

C. For Optimization of Gene Clustering Process

Sheng. et. al [7] proposes a niching genetic k-means algorithm (NGKA), to obtain the high quality clustering solution by minimizing the sum of square error criteria. The proposed algorithm focuses on the two aspects, firstly the niching method in which they combine those clusters solution showing similar property with the GAs which helps in removing the problem of premature convergence involved during the clustering search process. Secondly, they combine the k-means algorithm to the result obtained from the first step in order to improve the computational efficiency. Instead of replacing the biological inspired operators, the one step k-means operation is included in the regeneration step of the process. The proposed algorithm is compared with previous work. Experimental results show that the proposed niching method and the k-means hybridization is highly effective for the proposed GA based clustering and it outperforms two previously GA based clustering method consider for experimentation.

One of the problems associated with GA based clustering algorithm is the problem of “Degeneracy”. A Degeneracy problem occurs when same solution is repeated by multiple chromosomes which results in exploration of the same clusters repeatedly. As a result the search space will not explored efficiently. So Chang Dong-Xia. et. al [8] proposed a genetic algorithm with gene rearrangement i.e. (GAGR) which enhance the performance of the cluster .In the proposed algorithm the degeneracy of the chromosome is effectively removed results in the fast convergence of the evolution process. For building the path between two parent chromosome a new path based crossover is also introduced and in order to remove the regeneracy problem a gene rearrangement of the chromosome has been applied. These processes help in exploring the search space effectively. The GAGR algorithm outperforms in comparison to previously GA based K-means clustering algorithm.

Mukhopadhyay *et al.* [9] has proposed a Genetic Algorithm based fuzzy clustering method (GOGA) which include the concept of Gene Ontology (GO). The proposed algorithm combines the expression based and GO-based gene dissimilarity measure for the clustering of microarray gene expression data .The Gene Ontology (GO) [10] is one of the most popular and important ontologies which is used to know whether gene in the cluster share the same biological information or not. Two GO-based semantic similarity measures, namely Resnik's [11] and Weighted Jaccard [12] similarity measures have been used to find the dissimilarities between genes from GO point of view. Experiments conducted on the Yeast cell cycle data set indicate that the combination of expression-based and GO-based dissimilarities improved the performance of the cluster as compared to only expression-based clustering.

III. PSO BASED GENE CLUSTERING

SI algorithms such as PSO are mainly stochastic search and optimization techniques, based on the principles of collective behavior and self organization of insect swarms. They are efficient, adaptive and robust search methods producing near optimal solutions and have a large amount of implicit parallelism (i .e. all particles in problem space parallel update the value of the position and velocity vector). On the other hand, several tasks in bioinformatic involve optimization of different criteria (like energy, alignment score, overlap strength and so on); thereby making the application of SI tools (like PSO) more obvious and appropriate. The problems of bioinformatic seldom need the exact optimum solution, rather what they need are robust, fast and near optimal solutions, which SI algorithms like PSO are known to produce efficiently. Thus researchers applied PSO and its variants to solve the problem of Gene Clustering. The next section gives a short introduction of PSO followed by the description of how hybridization of PSO with K-means, GA and other metaheuristics has been applied by several researchers to solve the Gene Clustering problem efficiently.

A. Introduction

The Particle Swarm Optimization, proposed by James Kennedy and Russel Eberhardt in [13] is an algorithm for creating swarms of particle which moves in the search space i.e. Problem space searching for their goal (based on problem definition). The concept is stated as, each particle updates its position and velocity based on its own best value and best value in the whole swarm. For a swarm of n particles the ith particle is represented by a position denoted as $xi = (xi1, xi2, \ldots, xin)$ where n is the number of particles. Except the position, each particle of a swarm is represented in D dimensional space with a velocity $vi = (vi1, vi2, \ldots, vid)$. The particles explore in the search space with a velocity that is dynamically adjusted according to its own and neighbor’s performances. The standard PSO method updates the velocity and position of each particle according to the equations given below.

$$\begin{aligned} vid(t+1) = \omega\, vid(t) + c1\ \mathrm{rand}()\ (pid - xid) \\ + c2\ \mathrm{rand}()\ (pgd - xid) \end{aligned} \quad (1)$$

$$xid\ (t+1) = vid(t+1) + xid\ (t) \quad (2)$$

where $c1$ and $c2$ are two positive acceleration constants, rand () is a uniform random number in (0, 1), pid and pgd are the best positions found so far by the ith particle and all the particles respectively, t is the iteration count and ω is an inertia weight which is usually, linearly decreasing during the iterations. In Eq. (1), there are three parts: the first one shows the current speed of particle i.e. The present state, the second term is known as the cognition term which shows the thought of the particle itself a the last and third term is a social term that shows the ability of information sharing among the swarms.

B. Hybridization of PSO with K-means

A classic K-Means algorithm [14] is one of popular method to cluster gene expression data. But the main drawback of K-means is that it is easily gotten trapped in local optima. In order to improve the performance of K-Means, Zhihua. et. al [15] proposed a modified form of K-means algorithm known as PK-Means which is the combination of the Particle Pair Optimizer and K-Means clustering algorithm. The Particle Pair Optimizer is a variation on the traditional particle swarm optimization algorithm, which is a stochastic particle pair based optimization technique. The results of the PK-Means algorithms outperforms K-Means due to a factor of exploration added into it. Experimental results show that PK-means has a faster convergence rate and low computational load and also outperforms the previously known algorithms K-means, Fuzzy K-means respectively.

Yau-King Lam *et al.* [16] proposes an improved PK means clustering algorithm by adding a concept of clustering matching. Each cluster centroid in a particle is matched with its corresponding centroid in the best particle based on the closest distance. The cluster matching method is a two step process. In the step I the sequence of the cluster contained in the particle position is matched with the cluster contain in the position of the particle’s global best position on the basis of nearest distance by constructing a nearest matrix. In this nearest matrix the column has the entry of the clusters in the current particle position and the row has the entry of the cluster in the global best position of the particle respectively. After this the sequence of the cluster contained in the current particle position is rearranged according to the matching results. Experimental results demonstrated that the PSO-KM with the proposed scheme can outperform its original version and other methods in terms of compaction.

Yau King Lam *et al.* [17] proposes a novel algorithm XK-means which hybridize PSO with K-means so as to integrate the concept of exploration with exploitation. This is done by using K-means and an exploratory vector. The exploratory vector is added to each centroid before a K-means iteration. The algorithm includes two parameters namely ai and bi respectively. These parameters are problem dependent. The value parameter bi is set in such a way that it decrease monotonically. As a result the search space gets explored and the exploitation level gets increased. The proposed method is compared with the K-Means algorithm and the PK-Means algorithm. The following observations are reported. Firstly, the cluster compactness (measured by MSE value) and the stability (measured with MSE standard deviation) show better results than those obtained from the K-means and the PK-means algorithms. Secondly, the proposed method is at least three times faster than the latter algorithm and finally, the complexity of the proposed algorithm is low in comparison to K-means and PK-means algorithm.

Zhen *et al.* [18] proposes a Particle Swarm Optimizer (PSO) based memetic algorithm (MA) [19] and named it as memetic K-means algorithm (MKMA). The aim of proposed algorithm is to minimize the sum of squared distance which is one of the problem associated with the K-means algorithm. MA combines both the global and local search. The global search in MKMA is done with the help of one of the variants of PSO i.e. Comprehensive learning particle swarm optimizer (CLPSO) [20]. CLPSO maintains the diversity of the swarm which help in improving the exploration efficiency of the global search. In each iteration of CLPSO, the particle swarm is partitioned into a leader and a populace group based on fitness value. Among the particle swarm one third of the best particles are considered as leader particles and the remaining particles forms the populace group. They conduct experiments on two gene expression datasets and reveal that MKMA has consistently attained better performance in comparison with than of K-means, Fuzzy K-means and PK-means.

C. Hybridization of PSO with Genetic Algorithm

Enrique. et. al [21] compare the use of a Particle Swarm Optimization (PSO) and a Genetic Algorithm (GA) (both combine the Support Vector Machines SVM) for classification the high dimensional Microarray Data. Both the algorithms are used for finding samples of informative genes from thousands of genes. An SVM classifier with 10-fold cross-validation is applied in order to validate and evaluate the provided solutions. A PSO SVM algorithm finds interesting genes. However, a new version of PSO, called Geometric PSO, is evaluated for the first time in this work using a binary representation in Hamming space. Both approaches i.e. PSO SVM and GA SVM were tested on six cancer datasets and outcome of the proposed algorithm helps a lot in identifying the specific genes. Experiment results reveals 100% classification rate and few genes per subset (3 and 4) in most executions.

Since clustering gene expression data based on the similarity measure is becoming one of the important research areas in the field of Bioinformatics, Shutao. et. al [22] proposed a gene selection method using combined PSO with GA using SVM. First, the Wilcoxon rank sum test (Deng et al. [23]) is used to pre process the original gene expression data and then the proposed hybrid PSO/GA is adopted to select the most important gene subsets using tenfold cross validation (CV) scheme. Finally, a classifier is trained based on the gene subset obtained from PSO/GA which is used to predict the testing samples. The algorithm test on the three gene expression datasets namely Leukemia, Colon and breast cancer Data. The experimental results show improvement in the formation of gene cluster and also leads to better classification accuracy.

D. Hybridization of PSO with Others

Xiao et al. [24] proposed a hybrid approach based on Self Organizing Maps (SOM) and PSO to cluster the genetic data. The proposed hybrid SOM/PSO algorithm work in two stages. In the first stage of the proposed algorithm, SOM is used to cluster the dataset. In this stage either regular SOM or SOM with a conscience can be used. The SOM normally runs for 100 iterations and generates a group of weights. In the second stage, PSO is initialized with the weights produced by SOM in the first stage. Then PSO is used to refine the clustering process. The experimental result shows that the proposed algorithm outperforms the average quantization error and the typographic error as compared to SOM.

Although the PSO algorithm converges fast, but lacks in global convergence and may get trapped in local optima. So Sun et.al [25] proposed the Quantum-behaved Particle Swarm Optimization (QPSO) algorithm. Although QPSO possesses better global convergence behavior than PSO, but it also suffers from premature convergence. In a QPSO system, Since gbest position guide the whole swarm particles. Once the gbest position get converges to local optima the particle get misguided by this gbest and result in bad exploration of the problem space. To overcome the problem of QPSO Jun Sun et.al [25] introduced the revised version of QPSO by adding a Multi-Elistic strategy for Quantum-behaved Particle Swarm Optimization (QPSO) known as Multi-Elitist QPSO model, which is drawn from Swagatam et.al [26], to update the gbest position of the QPSO algorithm. In the MEQPSO algorithm , the particle's search is not based on the global position i.e. gbest but on the bases of the promising search region so that the particle has more chance to search in this region and obtain the global optimal solution. As a result , the MEQPSO have a stronger global search ability and better overall performance that the original QPSO.

IV. CONCLUSION

The bioinformatics field requires the advanced computing tools for processing large amounts of biological data and involve various problems such as Gene Clustering Problem, Molecular Docking Problem, Multiple Sequence Alignment Problems, Phylogenetic tree construction , RNA Secondary Structure Prediction etc.and among them gene clustering is investigated in this paper.Gene Clustering becomes popular nowadays because of matured microarray technology and increasing power of computing Although there are many clustering algorithms developed, K-means is one of the popular methods for gene clustering due to its effectiveness and simplicity. But it suffers from the disadvantage of choosing the initial centroid and easily get trapped in local optima. So to enhance its efficiency population based algorithms, such as PSO and GA is incorporated with K-means clustering by several researchers as described in this paper. Also from the survey it has been found that main challenge in gene expression analysis to determine the relevant gene expression data which is of high dimension , redundant, noisy and having a small sample size. Gene selection can improve the prediction accuracy of classifiers and save computational cost by using only discriminating genes with reduced dimensionality. Another problem while analyzing gene expression data occurs when the dataset involve multiple classes. So, the classical deterministic search algorithm is of no use for them as the search space may be very large and discontinuous at several points. Thus population based algorithm such as PSO and GA, have recently gained wide popularity among the researchers, for their ability to find robust and near optimal solution for such an NP hard problem. GA and PSO both suffer from the problem of local optima. So their integration with each other and with other metaheuristic were done by several researchers. This article surveyed work done regarding the hybridization of population based algorithm particularly GA and PSO with others for solving Gene Clustering problem.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their detailed, valuable comments and constructive suggestions.

REFERENCES

[1] L. Li, T. A. Darden, C. R. Weinberg *et al*., "Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method," *CombChemHigh Throughput Screen* vol. 4 , pp.727–739, 2001.

[2] C. H. Ooi and P. Tan, "Genetic algorithms applied to multi-class prediction for the analysis of gene expression data," *Bioinformatics* vol. 19, pp. 37–44, 2003.

[3] S. Penga, Q. Xub, X. B. Lingc, X. Pengd, W. Dua, and L. Chenb, "Molecular classification of cancer types from microarray data using the combination of genetic algorithms and support vector machines," *Elsevier B.V. on behalf of the Federation of European Biochemical Societies*, pp. 358–362, 2003.

[4] V. N. Vapnik, “Statistical learning theory,” Wiley, New York 1998.

[5] S. Ramaswamy, *et al.*, Natl. Acad. Sci. USA 98 , pp. 15149-15154, 2001.

[6] C.-P. Lee,W.-S. Lin,Y.-M. Chen, and B.-J. Kuo, “Gene selection and sample classification on microarray data based on adaptive genetic algorithm/k-nearest neighbor method,” *Expert Systems with Applications* vol. 38, pp. 4661–4667, 2011.

[7] W. Sheng, A. Tucker, and X. Liu, “A niching genetic k-means algorithm and its applications to gene expression data,” *Springer-Verlag*, pp. 9-19, 2008.

[8] D.-X. Chang, X.-D. Zhang, and C.-W. Zheng, “A genetic algorithm with gene rearrangement for K-means clustering,” *Pattern Recognition*, pp. 1210-1222, 2009.

[9] A. Mukhopadhyay, U. Maulikt, S. Bandyopadhyay, and B. Brors, “GOGA: GO-driven Genetic algorithm-based fuzzy clustering of gene expression Data,” in *Proc. of 2010 International Conference on Systems in Medicine and Biology*, 16-18 December 2010, IIT Kharagpur, India, pp. 221-226.

[10] M. Ashburner, C. A. Ball, J. A. Blake, *et al.*,"Gene Ontology: tool for the unification of biology. The gene ontology consortium," *Nature genetics*, vol. 2s, no. I, pp. 25-29, 2000.

[11] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," in *Proc. of 14th international joint conference on Artificial intelligence (JJCAf-95)*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., pp. 448-453, 1995.

[12] C. Pesquita, D. Faria, H. Bastos, A. O. Falco, and F. M. Couto, “An information-theoretic definition of similarity," in *Proc. of* Lath Annual BioOntologies Meeting, pp. 37-40, 2007.

[13] J. Kennedy and R. C. Eberhardt, “Particle swarm optimization,” in *Proc. of IEEE International Joint Conference on Neural Networks, IEEE*, Piscataway, pp.1942–1947,1995.

[14] J. A. Hartigan and M. A. Wong, “A k-Means clustering algorithm,” *Appl. Stat.*, vol. 28, pp. 126–130, 1979 .

[15] Z. Du, Y. Wang, and Z. Ji, “PK-means: A new algorithm for gene clustering,” *Computational Biology and Chemistry*, pp. 243-247, 2008.

[16] Y.-K. Lam, P. W. M. Tsang, and C.-S. Leung, “Improved gene clustering based on particle swarm optimization, K-Means, and cluster matching.,” *Springer-Verlag Berlin Heidelberg*, pp.654-661, 2011.

[17] Y.-K. Lam, P. W. M. Tsang, and C.-S. Leung, “eXploratory K-Means: A new simple and efficient algorithm for gene clustering,” *Applied soft computing*, pp. 1149-1157, 2012.

[18] Z. Ji, W. Liu, Z. Zh, “Gene Cluster in using Particle Swarm Optimizer based Memetic Algorithm,” Springer-Verlag Berlin Heidelberg, 2011.

[19] P. Moscato, “Memetic Algorithm: a Short Introduction,” McGraw-Hill, London 1999.

[20] J. J. Liang and A. K. Qin *et al.*, ”Comprehensive learning particle swarm optimizer for global optimization of multimodal functions,” in *Proc. IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 81–295, 2006.

[21] E. Alba, Jos´e Garc´ıa-Nieto, L. Jourdan, and El-G. Talbit, “Gene selection in cancer classification using PSO/SVM and GA/SVM Hybrid Algorithms,” Congress on Evolutionary Computation, Singapore : Singapore 2007.

[22] S. Li, X. Wu, M. Tan, “Gene selection using hybrid particle swarm optimization and genetic algorithm,” Springer-Verlag, pp. 1039–1048, 2007.

[23] L. Deng, J. Pei, J. Ma *et al.*, “A rank sum test method for informative gene discovery,” *The 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 410–490, 2004.

[24] X. Xiao, E. R. Dow, R. Eberhart, Z. B. Miled, and R. J. Oppelt, *Gene Clustering using self-organizing maps and particle swarm optimization*, 2003.

[25] J. Sun, W. C. Fang, X. Wun, and W. Xu, “Gene expression data analysis with the clustering method based on an improved quantum-behaved Particle Swarm Optimization,” *Engineering Application of Artificial Intelligence*, pp. 376-391, 2012.

[26] D. Swagatham, A. Ajith, and K. Amit, “Automatic kernel clustering with a materialistic particle swarm optimization algorithm,” *Pattern Recognition Lett.*, vol. 29, pp. 688-699, 2008.

Arpit jain was born in Jabalpur (M.P) India on 20th Oct 1987. He receives the B.E. degree in Computer Science from Rajiv Gandhi Proudyogiki Vishwavidalaya Bhopal in 2009. He is presently persuing MTech from Rajiv Gandhi Proudyogiki Vishwavidalaya Bhopal.His area of research is Particle Swarm Optimization, Bioinformatics and Data Mining.

Shikha Agrawal is an Assistant Professor in Department of Computer Science & Engineering at University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M P) India. She obtained B.E., M.Tech., She has ten year of teaching experience. She is working on Swarm Intelligence for Multimodal function optimization. Her area of interest is Artificial Intelligence, Soft Computing and Particle Swarm Optimization. She has published more than 10 research papers in different reputed national and internal conferences and journals

Dr. Jitendra Agrawal was born in 1974, obtained his master degree in computer technology in 2005 and Ph.D. degree in Information Technology in 2012. Currently Dr Agrawal is working as a Assistant Professor in School of Information Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M. P.) India. He has 15 years of teaching experience and actively involved in research with area of interest as Soft Computing and Data Mining. He has published more than 20 research papers in national and international journals. He has more than 22 research publications to his credit in national and international journals of repute.

Dr. Sanjeev Sharma has graduated in Electrical & Electronics from Samrat Ashok Technical Institute, India and post graduated in Microwave and Millimeter from Maulana Azad College of Technology, India. He completed his Doctorate in Information Technology from Rajiv Gandhi Proudyogiki Vishwavidyalaya. Currently he is working as head in School of IT, RGPV (Bhopal), India. He possesses teaching and research experience of more than 17 years. His areas of interest are Mobile Computing, Data Mining and Information Security. He has edited proceedings of several national and international conferences and published more than 70 research papers in reputed journals.