Volume 4, Issue 6, June 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Preventing Premature Convergence in Genetic Algorithm Using DGCA and Elitist Technique

Ms. Shikha Malik[1], Mr. Sumit Wadhwa[2]
[1]M.Tech Student, CSE Department, [2]Assistant Professor, CSE Department
[1,2]Samalkha Group of Institution, Kurukshetra University, India

Abstract— ***Classic GA has normally a problem i.e. premature convergence problem which occurs when genes of high rated individuals quickly attain to dominate the population, constraining it to converge to a local optimum. This phenomenon occurs when the population of a genetic algorithm reaches a suboptimal state that the genetic operators can no longer produce offspring with a better performance than their parents. To avoid the premature convergence we have to preserve the population diversity during the evolution. Proposed algorithm will show improved results in terms of both computational time and quality of solution. The use of both mutation and crossover operators make them highly immune to be trapped in local optima and thus less vulnerable to premature convergence problem. In this work the concept of DGCA and elitist technique have been proposed and applied to travelling salesman problem (TSP) to reduce the effect of premature convergence problem. This will rapidly increase the performance of GA because this prevents losing the best found solutions. To implement the proposed methodology MATLAB R2009a has been used as a software platform. Our primary objective is to increase the efficiency of result set with discarding the worst solution by maintain diversity of structures in the population.***

Keywords— Genetic algorithm, Premature convergence, DGCA, Elitist, Diversity

I. INTRODUCTION

Genetic Algorithm (GA) is adaptive heuristic based on ideas of natural selection and genetics. Genetic algorithm is one of the most known categories of evolutionary algorithm. A GA works with a number of solutions which collectively is known as population in each iteration which is chosen randomly. These are adaptive heuristic search algorithms postulated on the evolutionary ideas of natural selection and genetic. The basic concept of these evolutionary algorithms is to stimulate process in natural system necessary for evolution. GA's are used for numerical and computational optimization and based on study the evolutionary aspects of models of social systems. The GA performs a balanced search on various nodes and there is a need to retain population diversity exploration so that any important information cannot be lost because there is a great need to focus on fit portions of the population. The basic GA operators are Crossover, mutation, selection. GA's suffer from the difficulty of local optimum convergence. It is the case when an extraordinary individual take over significant proportion of the finite population and leads towards the undesirable convergence. This problem is named as premature convergence or diversity. There is a great need to prevent premature convergence or diversity.

A. Premature Convergence

The premature convergence of a genetic algorithm arises when the genes of some high rated individuals quickly attain to dominate the population, constraining it to converge to a local optimum. The premature convergence is generally due to the loss of diversity within the population. This loss can be caused by the selection pressure, the schemata distribution due to crossover operators, and a poor evolution parameters setting. This phenomenon occurs when the population of a genetic algorithm reaches a suboptimal state that the genetic operators can no longer produce offspring with a better performance than their parents. To avoid the premature convergence, in a genetic algorithm is imperative to preserve the population diversity during the evolution.

B. Travelling Salesman Problem

Proposed technique will be applied to travelling salesman problem (TSP) to reduce the effect of premature convergence problem. Travelling salesman problem (TSP) is an NP-hard problem in combinatorial optimization technique. This problem deals in finding the shortest tour that passes exactly once through each vertex in a given graph. The idea behind the TSP is to find a tour of a given number of cities, by visiting each city exactly once and return to the starting city where the length of this tour is minimized [15]. This will try to cover as many locations as possible without visiting any location twice is the most important aspect of the scheduling of a tour.

The travelling salesman problem consists of a salesman and a set of different number of cities. The salesman has to visit each city starting from a specific one (e.g. the hometown) and returning to the same city. The challenge of the problem is that the travelling salesman wants to minimize the total distance of the trip.

II. RELATED WORK

Bahareh Nakisa *et. al.* [1] presented a comprehensive survey of the various PSO-based algorithms such that PSO is a computational search and optimization method based on the social behaviours of birds flocking or fish schooling. A number of basic variations have been developed due to solve the premature convergence problem and improve quality of solution founded by the PSO. As part of this survey, the author include a classification of the approaches and we identify the main features of each proposal. P.K.Suri *et. al.* [2] proposed that Elitism technique was augmented within Genetic Algorithm allowing the best solution from any generation to be carried across the new population allowing it to sustain. Social Disaster Techniques (SDT) were used when premature convergence occurred and the problem of premature convergence may be avoided by creating random offspring and inserting diversity in the population. This paper attempted to use the both concepts of Elitism and Social Disaster techniques spanning across various generations. A previous solution was chosen and it has been looked upon how Elitism and Social Disaster techniques fares towards the same problem. Deepti Gupta *et. al.* [5] proposed that Genetic algorithm is a search & optimization method based on the Darwin's principle of Survival of the fittest. It is an abstraction of complex natural genetics and natural selection process. Genetic algorithm is based on the principle of natural selection for reproduction and various evolutionary operations as crossover and mutation. Two controlling factors that need to be balanced in the process of selection are Genetic Diversity and Selective Pressure. Population Diversity could be controlled by a means of ways as Fitness sharing, Deterministic crowding and so many other. In this paper the author was providing a brief knowledge about variety of methods maintaining population diversity. Rakesh Kumar *et. al.* [6] described that in roulette wheel selection operator there is essence of exploitation while rank selection is influenced by exploration. In this paper, a blend of these two selection operators is proposed that is a perfect mix of both i.e. exploration and exploitation. The blended selection operator is more exploratory in nature in initial iterations and with the passage of time, it gradually shifts towards exploitation. The proposed solution is implemented in MATLAB using travelling salesman problem and the results were compared with roulette wheel selection and rank selection with different problem sizes. Chaiwat Jassadapakorn *et. al.* [8] considered a situation where the evolutionary process converges too fast to a solution which causes it to be trapped in local optima. To overcome this problem, a proper diversity in the candidate solutions must be determined. Most existing diversity maintenance mechanisms required a problem specific knowledge to setup parameters properly. This work proposed a method to control diversity of the population without explicit parameter setting. A self-adaptation mechanism was proposed based on the competition of preference characteristic in mating. It can adapt the population toward proper diversity for the problems. The performance of the adaptive method is comparable to traditional Genetic Algorithm with the best parameter setting. Er.Rajiv Kumar *et. al.* [9] assume that to maximize the performance of genetic algorithm there is a need to study the convergence state of genetic algorithm. Genetic algorithm is a Mata-heuristic search technique; this technique is based on the Darwin theory of Natural Selection. The important property of this algorithm was that it has worked on multiple state of solution. This algorithm worked with some finite set of population. Each member of the population is represented by a string written over fixed alphabets and also each member has a merit value associated with it, which represent its suitability for the problem under consideration. In this paper the author study the effect of crossover and inversion probability on the convergence of genetic algorithm. The convergence of genetic algorithm is depends upon the parameter setting of genetic algorithm. Elena Simona Nicoara *et. al.* [10] described that the optimization by genetic algorithms often comes along with premature convergence bias, especially in the multimodal problems. In the paper the author propose and test two mechanisms to avoid the premature convergence of genetic algorithms by preserving the population diversity in two different manners. These were the dynamic application of many genetic operators, based on the average progress and the population partial re-initialization. The mechanisms were tested by implementing them in the NSGA_II algorithm, applied to one of the most difficult job shop scheduling test problems, ft10. The comparative analysis between the new algorithm and the NSGA_II in the absence of the submitted mechanisms alongside with an elitist and the canonic genetic algorithm proved the usability of both proposed mechanisms. Guilan Hu [11] suggested that personal information mining can find out the hidden relationship and characteristics of the target people which can be used for active post operation. The original features which were included in the personal information raw data usually have high dimension and redundancy which often drags down data mining efficiency. A feature optimization method was proposed to resolve the problem. The method with the purpose of data dimensionality deduction was based on effective association of rough set theory with PCA approach. The final classification features were derived through two steps of optimization and deduction operation. Some documents data from education field and financial field were used for the experiment. The experimental results demonstrate that the hybrid feature optimization method was effective in improving classification accuracy. Sofiane Achiche *et. al* [14] suggested that the growing number of industrial fields were concerned by complex and multiobjective problems. The main drawback was the need of an expert to manually construct the knowledge base. The use of genetic algorithms proved to be an effective way to solve this problem. A main problem in genetic algorithms was the premature convergence, and the last enhancements in order to solve this problem include new multi-combinative reproduction techniques. There were two principal ways to perform multi-combinative reproduction within a genetic algorithm, namely the Multi-parent Recombination, Multiple Crossover on Multiple Parents (MCMP); and the Multiple Crossovers per Couple (MCPC). Both techniques try to take the most of the genetic information contained in the parents. This paper explored the possibility to decrease premature convergence in a real/binary like coded genetic algorithm (RBCGA) used in automatic generation of fuzzy knowledge bases (FKBs). The RBCGA used several crossover mechanisms applied to the same couple of parents. The crossovers were also combined in different ways creating a multiple offspring from the same parent genes. The large family concept and the variation of the crossovers should introduce diversity and variation in otherwise prematurely converged populations and hence keeping the search process active.

III. PROPOSED MODEL

To implement the proposed methodology MATLAB R2009a has been used as a software platform. In earlier models two approaches were used such as the dynamic application of crossover and mutation operators and the population partial re-initialization. The proposed model comprise of two techniques such as Elitist and DGCA (Dynamic Genetic Clustering Algorithm).

Various steps of proposed model are as follows:

1. Randomly generate the population set of individuals.
2. Evaluate each individual by a fitness function.
3. Divide the complete population into clusters using DGCA technique.
   (Repeat step 4 to 8 for every cluster)
4. Now divide each cluster into two equal parts.
5. Select the best chromosomes from first half of cluster and second half of cluster.
6. Apply crossover operation on these two selected best chromosomes.
7. Select best one out of these chromosomes (i.e. best of first half, best of second half and child of these two after crossover operation) using elitist.
8. Mutate it in different ways hence we get chromosomes for new population from one best chromosome equal to the size of each cluster. Pass these chromosomes to next generation.
9. Hence we get the new population.
10. Repeat the whole process again till we get the optimum solution.

A. *Execution Model*

1) *Implementation with normal genetic algorithm:* This model implements with basic Genetic algorithm. Initially consider entire population and use tournament selection to get 2 fittest individuals and then perform crossover over them to get new chromosome. Then mutation is performed on that chromosome and result is forwarded to new population repeat again until the new population of same size occurred. In 1st iteration fittest solution is fetched then in second iteration again the fittest solution is fetched if the result of second iteration is best than first so, swapping is performed otherwise first result is considered fittest. Repeat this for say, 1000 or 2000 iterations finally the best path is obtained. This leads to the problem of premature convergence because tournament selection is performed on complete population so it tend to converge towards local optimum.

2)

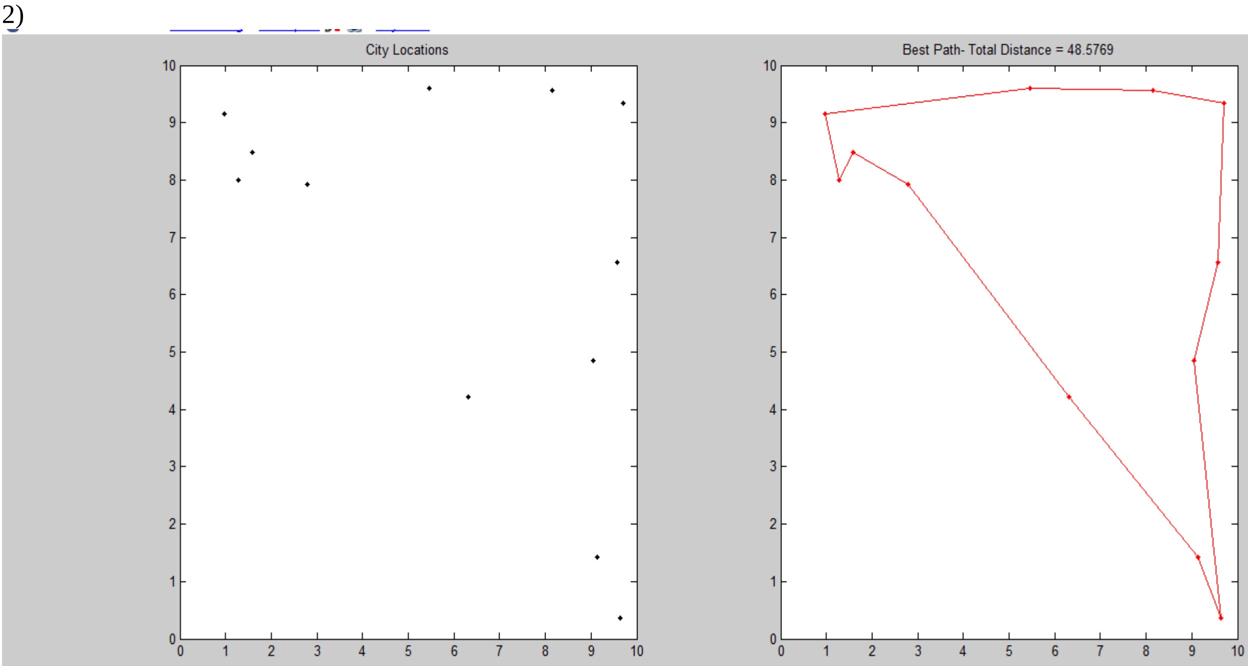


Fig 1 City locations with best path selected in normal GA

This figure has two graphs. First shows the randomly generated city locations represented by dots (.) and the second shows total distance of connected cities having premature convergence problem. The peak in upper left shows this problem.

3) *Implementation with DGCA and Elitist technique:* This model implements with DGCA and Elitist technique to avoid premature convergence and the results are compared with the results obtained from using simple GA. Initially DGCA technique is applied to create clusters among the population. Clustering means to divide complete population in clusters so as to avoid premature convergence. Number of clusters are calculated using formula-

x=round(log2(psize));

Where x is no. of clusters formed
and psize is population size
Calculate best fittest individual from each cluster. For this we have to divide every cluster into two parts and fittest individual is obtained from each cluster part and then crossover is performed on these fittest individuals.

For example- P1 is parent 1 which is fittest individual from first half of cluster and P2 is parent 2 which is fittest individual from second half of cluster. C1 is the child produced by crossover operation. Here crossover point is set to 6. So digits from 1 to 6 are copied from P1 and order of the sequence of digits after 6 are taken from P2.

↓ (crossover point)
P1 = |1 2 3 4 5 6| 7 8 9 10
P2 = 2 8 1 10 3 9 5 7 4 6

↑ ↑ ↑ ↑

(i) (ii) (iii) (iv)

C1 = 1 2 3 4 5 6 8 10 9 7

After that we select best out of these three individuals (two parent and one child) using Elitist technique. Elitist technique will choose the fittest individual whether it is a parent or a child. To create new population mutation is performed on the best fittest solution equal to the size of cluster. This will remove premature convergence.

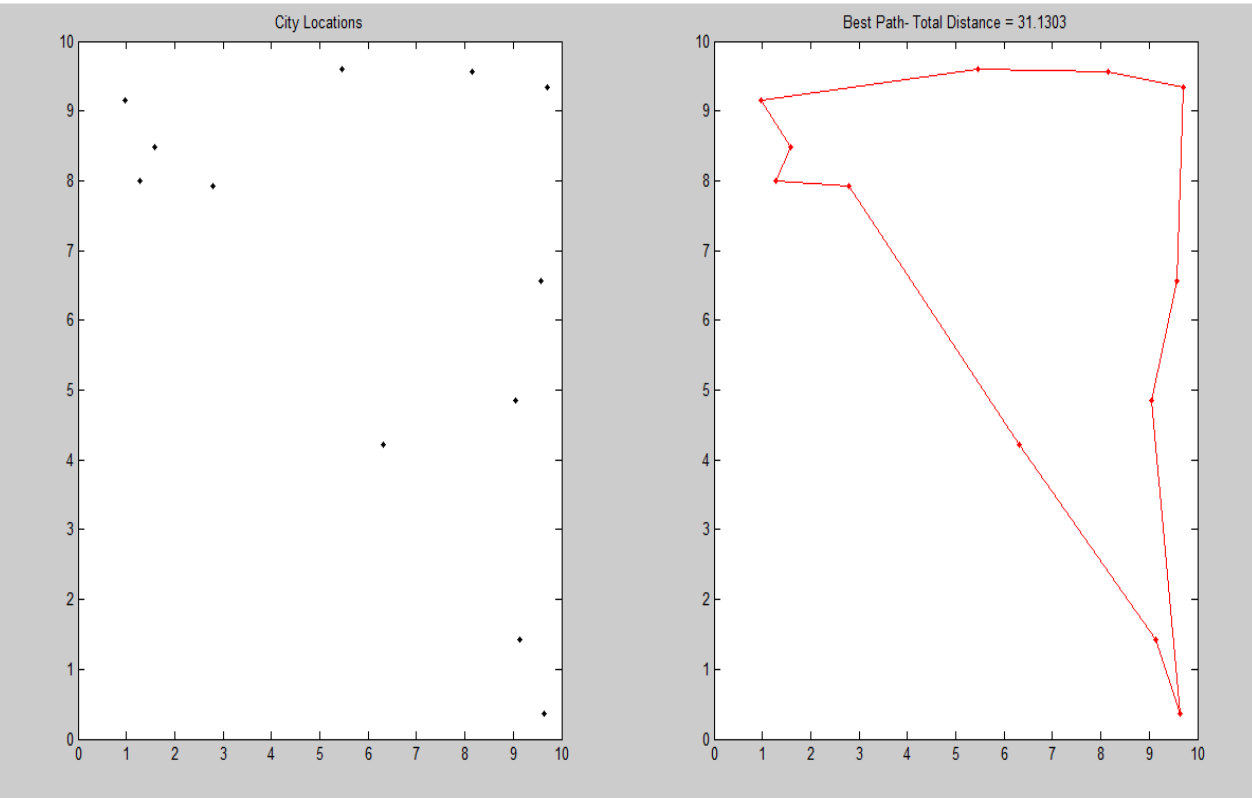
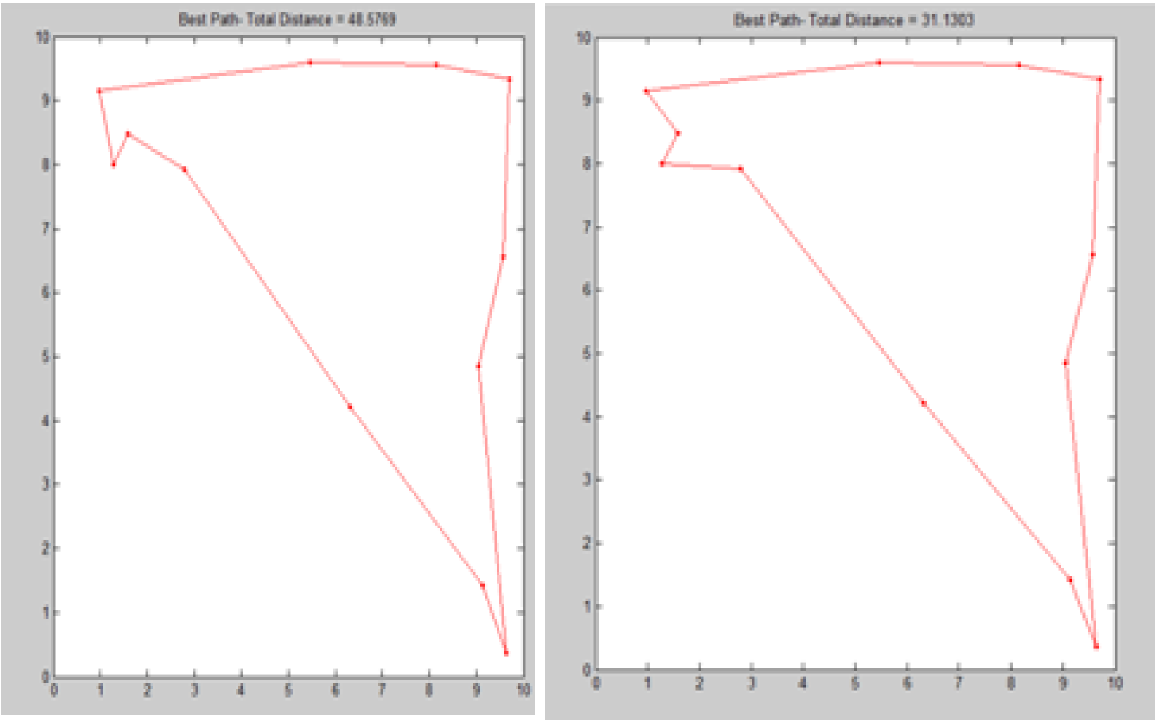


Fig 2 City locations with best path selected in proposed scheme

The above figure has two graphs. First shows the randomly generated city locations represented by dots (.) and the second shows connected cities so that the distance travelled by covering all the cities exactly once is minimum.

4) Comparison between normal GA and Proposed Technique: When we compare our proposed technique with normal GA it is very clear that proposed technique gives better results. The following figure shows this.



Normal GA	**Proposed technique**
No. of cities = 12	No. of cities = 12
Best Path = 48.5769	Best Path = 31.1303
no. of iteration= 436	no. of iteration= 94
(premature convergence problem here)	(No premature convergence problem here)

Fig 3 Comparison between normal GA and Proposed Technique

B. Simulation results

1) Results obtained using Normal genetic algorithm

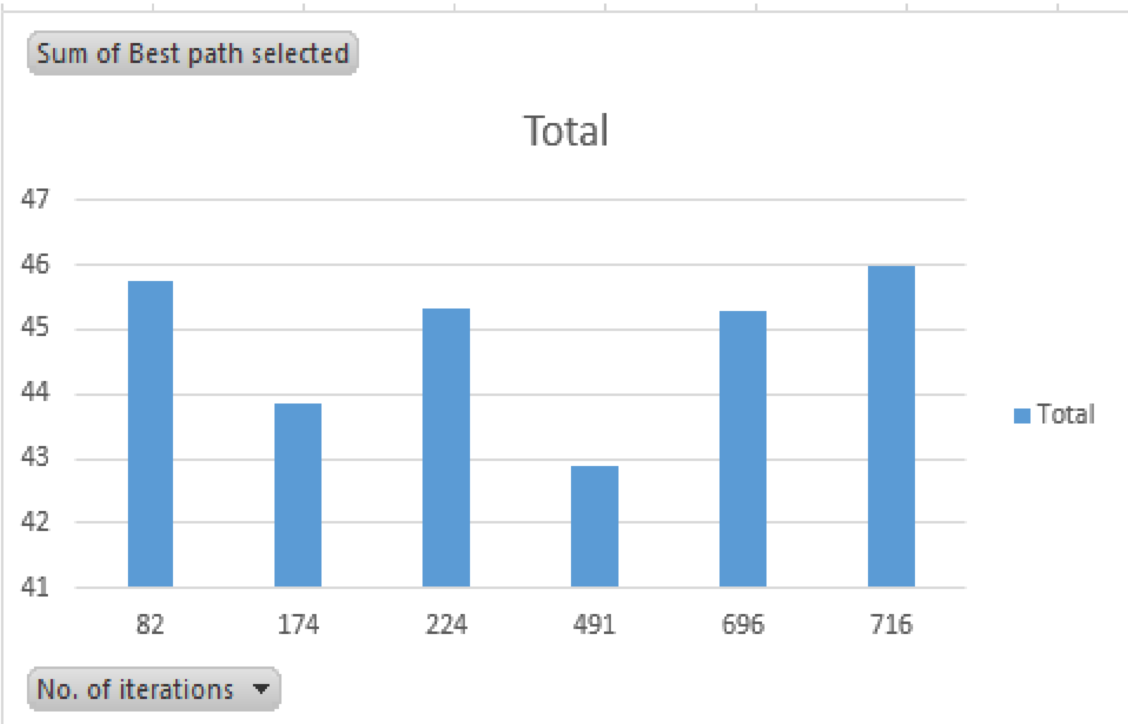


Fig 4 Graph between no. of iterations VS best path obtained using Simple GA

In this graph the number of cities were maintained fixed so as to calculate the best path with different number of iterations. This will show in how many iterations for same number of cities we get the best optimum path length.

TABLE I
SIMULATION RESULTS FOR NO. OF ITERATION VS BEST PATH USING SIMPLE GA

No. of Iterations	**Best Optimum path**
82	45.76
174	43.85
224	45.34
491	42.87
696	45.302
716	46

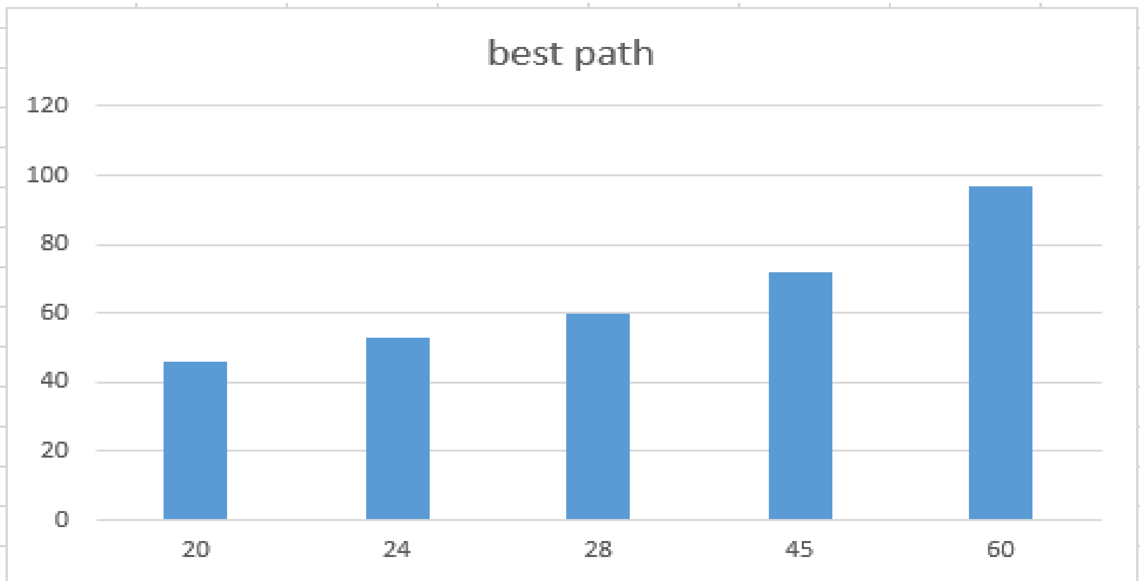


Fig 5 Graph between no. of cities VS best path obtained using Simple GA

In this graph at each time of execution number of cities were different which provide a different value for best path. This will provide an overview to check the best path with different number of city locations. The best path obtained is not the global optimum path because the normal GA has premature convergence problem. So the result produces are less efficient.

TABLE II
SIMULATION RESULTS FOR NO. OF CITIES VS BEST PATH USING SIMPLE GA

No. of cities	**Best Optimum path**
20	46
24	53
28	60
45	72
60	97

2) Results obtained using DGCA and Elitist technique

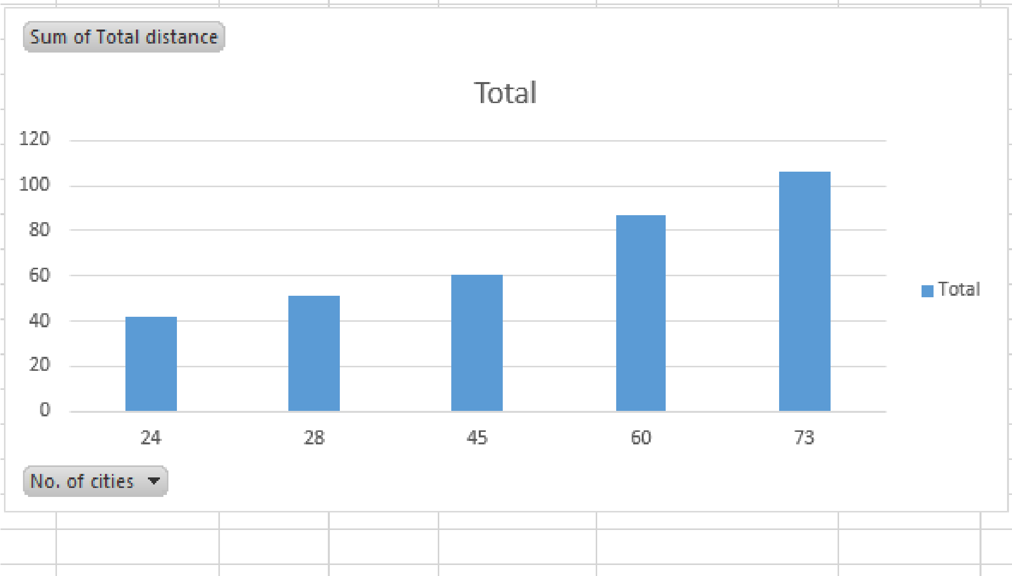


Fig 6 Graph between no. of cities VS best path obtained using Proposed Scheme

In this graph at each time of execution number of cities were different which provide a different value for best path. This will provide an overview to check the best path with different number of city locations. When these results are compared with the results obtained from normal genetic algorithm this shows with same number of cities the value of best optimum path is less.

TABLE III
SIMULATION RESULTS FOR NO. OF CITIES VS BEST PATH USING PROPOSED SCHEME

No. of cities	**Best Optimum path**
24	42.107
28	50.99
45	60.53
60	86.71
73	106.68

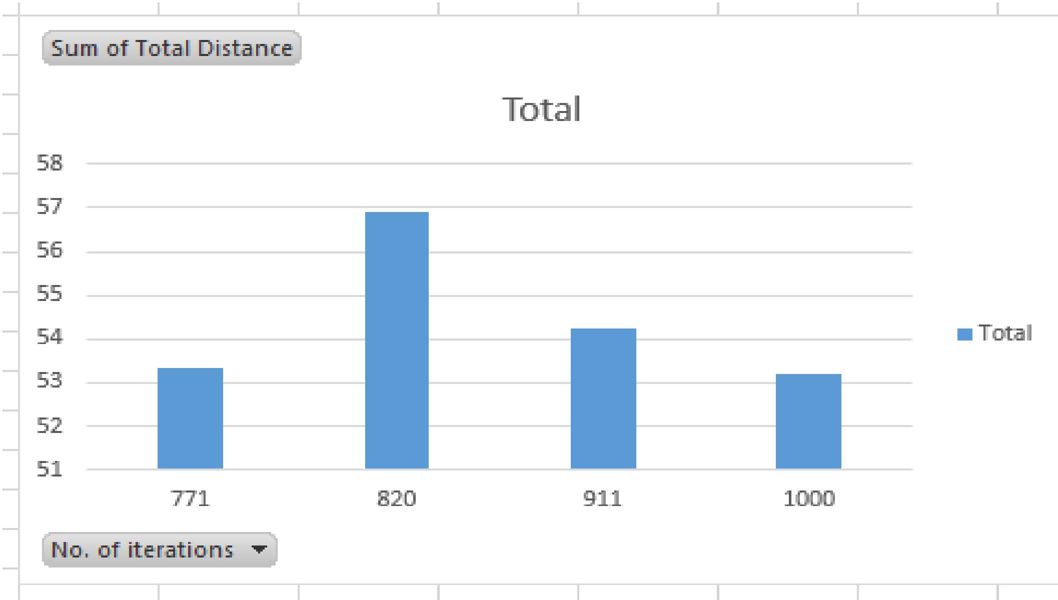


Fig 7 Graph between no. of iterations VS best path obtained using Proposed Scheme

In this graph the number of cities were maintained fixed so as to calculate the best path with different number of iterations. This will show in how many iterations for same number of cities we get the best optimum path length. When these results are compared with normal GA results, it shows that our proposed scheme uses less number of iterations to produce best optimum path. Hence computation time is lower than the previous technique. It can be concluded from these graphs that our proposed technique is more efficient and is very fast as compared to previous one i.e. normal GA.

TABLE IV
SIMULATION RESULTS FOR NO. OF ITERATIONS VS BEST PATH USING PROPOSED SCHEME

No. of Iterations	**Best Optimum path**
771	53.33
911	54.26
1000	53.18
820	56.90

IV. CONCLUSION AND FUTURE WORK

Although Genetic algorithm is a good technique to get optimized result but in case of global optimum solution it fails because of its main problem i.e. premature convergence. So as to overcome this problem hybrid technique is used which is composed of DGCA and Elitist method which involves each individual from the complete population and provide a best fittest solution. Using the proposed technique the best global optimum solution is obtained. The result shows both the diversity and stability is maintained. The diversity factor reduced the chance of premature convergence and therefore reduced the chance that GA will be trapped in a local optima while the stability factor stabilized the population so the proposed model can take advantage from the information gathered over generations among the chromosomes about the optimal solution. In future works this hybrid model can be implemented on various problems to maintain diversity and get optimum solution. To maintain the diversity and reduce premature convergence various other techniques were present which can be further implemented on this problem and the results can be compared so as to check the best technique.

REFERENCES

[1] Bahareh Nakisa, Mohd Zakree Ahmad Nazri, Mohammad Naim Rastgoo and Salwani Abdullah "A Survey: Particle Swarm Optimization Based Algorithms To Solve Premature Convergence Problem" Journal of Computer Science 10 (10): 1758-1765, 2014

[2] P.K. Suri, Rakesh Kumar, Pardeep Kumar Mittal "A Simulator to Prevent Premature Convergence in YM using GA" International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS), pp:- 99-106, 2013

[3] Nitin S. Choubey and Madan U. Kharat " Performance Evaluation of Methods for handling Premature Convergence in GA - Case of Grammar Induction" International Journal of Computer Applications (0975 – 8887) Volume 79 – No.2, pp:- 9-13, October 2013

[4] C. Vanaret, J.B. Gotteland, N. Durand and J.M. Alliot "Preventing Premature Convergence and Proving the Optimality in Evolutionary Algorithms" International Conference on Artificial Evolution, Bordeaux, version 1-7, pp:-84-94, 2013

[5] Deepti Gupta and Shabina Ghafir "An overview of methods maintaining diversity in genetic algorithms" International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2, Issue 5, May 2012

[6] Rakesh Kumar and Jyotishree "Blending Roulette Wheel Selection & Rank Selection in Genetic Algorithms" International Journal of Machine Learning and Computing, Vol. 2, No. 4, pp:- 365-370, August 2012

[7] Majid Yousefikhoshbakht, Ebrahim Mahmoodabadi and Mohammad Sedighpour "A modified elite ACO based avoiding premature convergence for traveling salesmen problem" J. Ind. Eng. Int., 7(15), pp: - 68-75, ISSN: 1735-5702, South Tehran Branch, 2011

[8] Chaiwat Jassadapakorn1 and Prabhas Chongstitvatana "Self-Adaptation Mechanism To Control The Diversity Of The Population In Genetic Algorithm" International Journal of Computer Science & Information Technology (IJCSIT) Vol. 3, No 4,pp:- 111-137, August 2011

[9] Er. Rajiv Kumar and Er. Ashwani Kaushik "Premature Convergence And Genetic Algorithm Under Operating System Process Scheduling Problem" Journal Of Global Research In Computer Science, Volume 1, No. 5,Pp: - 7-11, December 2010,

[10] Elena Simona Nicoara "Mechanisms to Avoid the Premature Convergence of Genetic Algorithms" Vol. LXI No. 1, pp:- 87 – 96, 2009

[11] Guilan Hu "Research on Feature Optimization Method in Personal Information Mining" International Conference on Information Engineering, 2009

[12] Gregory S. Hornby "The Age Layered Population Structure for Reducing the Problem of Premature Convergence" ACM, July 2006

[13] Heni Ben Amor and Achim Rettinger "Intelligent Exploration for Genetic Algorithms" ACM, pp:- 25-29, July 2005

[14] Sofiane Achiche, Marek Balazinski and Luc Baron "Multi-Combinative Strategy To Avoid Premature Convergence In Genetically-Generated Fuzzy Knowledge Bases" Journal Of Theoretical And Applied Mechanics 42, 3, pp. 417-444, Warsaw 2004

[15] Chiung Moon, Jongsoo Kim, Gyunghyun Choi and Yoonho Seo "An efficient genetic algorithm for the traveling salesman problem with precedence constraints" European Journal of Operational Research 140, 606–617, 2002

[16] S.N.Sivanandam, S.N.Deepa, Introduction to Genetic Algorithms, Springer Berlin Heidelberg, New York (2008)

[17] Goldberg, D. E., *Genetic* Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, New York, NY, U.S.A. (1989)