Modified Genetic Algorithm for Maximum Clique Problem

Naresh Kumar[#1], Deepkiran Munjal[*2]

[#1]M Tech Scholar, BSAITM Faridabad, M.D.U. Rohtak
[*2]Assistant Professor, BSAITM Faridabad, M.D.U. Rohtak

Abstract— Maximum clique problem is NP-hard problem which applies its application in determining the maximum connected sub graphs. Cliques are one of the basic concepts of graph theory and are used in many other mathematical problems and constructions on graphs. Clique corresponds to individual set in the graph. The approach used in obtaining maximum clique is Genetic Algorithm. Genetic Algorithms are a type of optimization algorithm which is based on the theory of natural selection for generating solutions of the problem. Genetic Algorithms are adaptive heuristic search algorithms for the survival of the fittest. The Genetic Algorithm is a probabilistic search algorithm that iteratively transforms a set of mathematical objects of fixed length binary string. The maximum clique problem calls for finding the maximum sized sub graphs in a particular graph. The intent is to develop a method to find the optimal solution from huge set of solutions.

Keywords: *N-Puzzle Problem, NP hard, Genetic Algorithm, Optimization Algorithm.*

I. INTRODUCTION

Maximum clique problem finds the fully connected sub graphs from a given graph G= (V, E) in an arbitrary connected graph where V= {1, 2,….n} is set of vertex and E represents edges(x, y) set of G. The maximum clique problem is to find the maximum clique in a graph. The maximum clique problem is an NP hard problem i.e. P=NP. There is no polynomial time algorithm that can estimate the maximum clique. Genetic Algorithms are heuristics search algorithms to find the best possible solution from a given set of solutions so that optimization can be obtained. Genetic Algorithms are applied when exploration space is too large and the solution can be assigned some fitness value. The algorithm applied with many graphs provides satisfactory results. Before applying GAs to particular problem, a randomized algorithm is designed to work with graph theory to find the cliques with good performance. The paper has been organized as follows. It contains literature review and explanation of GAs. The work proposed has been explained and analysed in paper.

II. LITERATURE REVIEW

An extensive and comprehensive review proposes a sufficient way to maximum clique problem to find the gaps in the exiting technique and the proposed new technique. The papers of high quality journals have been analysed to get an idea of previous solution and as far as possible the review has been unbiased. In one of the work, randomly generated graphs with some heuristic functions were used for colouring of the graph. In another work, polynomial time algorithm using NP completeness by inherently intractable problems for determining cliques [9]. In other work, two algorithms were analysed for enumerating all maximal cliques [20]. One runs with o (M (n)) time delay and in o (n2) space and the other runs with o (4) time delay and in o (m+ n) space. In other work, an efficient Branch and Bound algorithm [16] for finding a maximum clique with computational experiment is described. An extensive review forms the foundation stone of a good research. Therefore an extra effort has been made in order to carry out the appraisal.

III. GENETIC ALGORITHM

A Genetic Algorithm is a parallel heuristic search procedure in which chromosomes and their genes represent solution of the problem. The Genetic Algorithm is inspired by search algorithms based on the mechanism of biological evolution. A genetic Algorithm is an iterative procedure maintaining a population of structures that are candidate solutions to specific domain challenges. Genetic Algorithm initiates the process of natural selection. Chromosomes are the initial element of GA. Chromosomes is made up of cells. The cells of finite length are represented by binary alphabet {0, 1}. The algorithm starts with a randomly generated population of available solutions. Genetic algorithms minimize the process of natural evolution to achieve optimization. For search problems genetic algorithms are used. Genetic algorithm requires encoding of parameter set so that fitness of each chromosome can be judged. The fitness value is a function for which chromosome is tested for its suitability. Genetic Algorithms should be used when alternate solutions are too slow or much complicated and want to hybridize with an existing solution. The basic element of GAs is problem encoding, selection method and Genetic operators such as crossover, mutation and replication.

1) Problem Encoding: Chromosomes and their genes represent solution of the problem. Genes have complex structure in a chromosome. The process of encoding generally depends on the properties of the problem statement. Binary encoding of bit string is performed in determining the cliques of a particular graph.

2) *Selection Method:* Selection methods are the basic criteria depend on fitness value decides to choose the individuals from the population. Selection methods also describe to replace the old population with the new population to avoid losing diversity of the problem. The purpose of selection method is to meet the requirement of the problems. The intention to apply selection is to prefer better solution than worst's. The chromosome with more fitness value will be considered better. Roulette wheel selection is used for selection from the population.

3) *Crossover:* extended to k-point crossover, where k crossover points are Crossover is generally performed to exchange the genetic material of two parents to produce new chromosomes. A crossover point is taken for which new population is generated. The point at which crossover is performed is depend on randomly selected crossover point. The number of crossover is calculated by the crossover rate which is generally 2-5%. The concept of crossover can be considered. Generally crossovers are of following types.

Single Point Crossover: In this crossover a single point is selected and applied on the two chromosomes. The binary string from the beginning is copied from one chromosome and the rest is copied from the other chromosome.

Parent 1 11010 01010
Parent 2 10101 10011
Child 1 11010 10011
Child 2 10101 01010

Two Point Crossover: Two crossover points are selected and crossover operator is applied. Binary string from beginning of chromosome to the first crossover point is copied from one parent, the part from the first to the second crossover point is copied from the second parent and the rest is copied from the first parent.

Parent 1 100110 1001 001

Parent 2 010110 0110 101

Child 1 100110 0110 001

Child 2 010110 1001 101

Uniform Crossover: In uniform crossover, bits are copied randomly from first or second parent uniformly.
10001011 + 11011101 = 10011111

4) *Mutation:* The Mutation operator guarantees the entire state-space will be searched, given enough time. Mutation causes movement in the search space locally or globally and involves reordering of the list. Mutation restores lost information or adds information to the population. It is performed on a child after crossover. Mutation is a genetic operator provides necessary condition to maintain genetic diversity from one generation of a population to the next generation of population. Mutation operator randomly flips bits in chromosomes. Mutation operator also termed as bit reversal or swap operator. Mutation can occur at each bit position in a string with some probability, generally very small (i.e., 0.001). The purpose of mutation in GAs is preserving and introducing diversity. Mutation provide the technique to avoid the local maxima by preserving the bit traversal from the population of the chromosomes.

Child 1 101 0 00010011

After mutation 101 1 00010011

5) Replication: The replication operator replaces the original parental population generated by selection, recombination and mutation.

IV. PROPOSED WORK

The work proposed to find the fitness value to achieve better optimization. The intent is to develop a method to find the optimal solution from huge set of solutions. Reordering is done on the basis of fitness value. Following steps are used to accomplish the proposed solution of the problem.

Step 1: An initial population of m chromosomes is generated of finite length of binary string.

Step 2: The chromosomes on n cells assigned on the vertices in the graph. To generate each cell of chromosome of 25 cells a random number is generated between 0 and 100 and if its value is greater than 50 the corresponding cell becomes 1 else it becomes 0.

Step 3: An index number is assigned for each of chromosomes. If the number of vertices is less and the sub graphs or clique are fully connected then the index will be higher.

Step 4: The fitness value is calculated for each chromosome according to formula $1/ (1+e - \lambda)$ where λ is vertex cover.

Step 5: A threshold value is described to define the range of index.

Step 6: Selection allocates number of copies to solution of problems having higher fitness value. Selection procedures include Roulette-Wheel selection, ranking selection etc.

Step 7: Crossover and mutation operators are applied for the regeneration of chromosomes based on the index number assigned to each chromosome.

Step 8: The population created by selection, crossover and mutation replaces the original population.

Step 9: Repeat steps until a terminating condition is met or optimization is achieved.

The following steps can be understood by diagram.

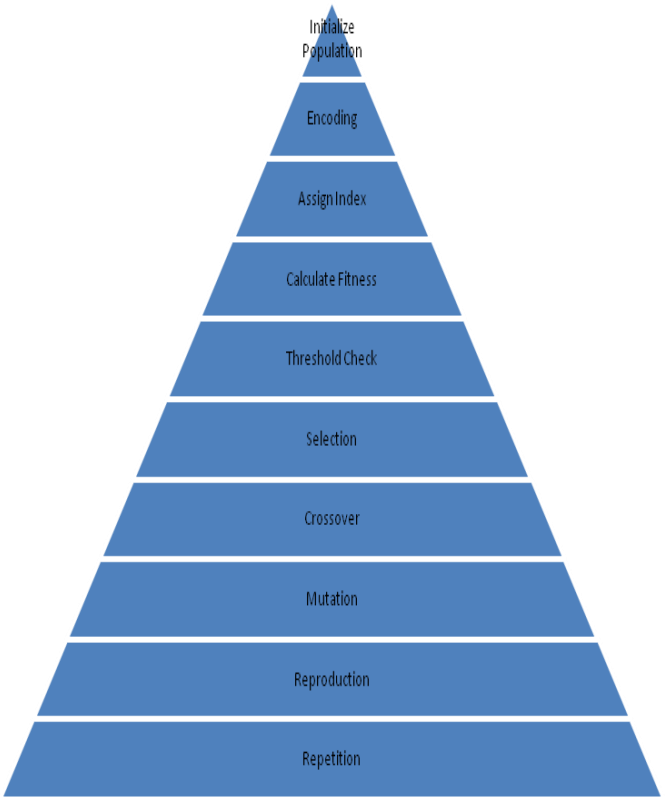

Fig. 1 A sample graph of Maximum Clique Problem

V. Result and Conclusion

In the above work, it has been established that applying the above technique results in generation of set of solution that provide approx certain answer to the problem. The problem statement therefore reaches to find out maximum vertex cover of a given graph using GAs. The method proposed is verified by varying parameters of Genetic Algorithm. The method proposed above presents a solution which has complexity far below the known methods. The algorithm has been tested limited number of time due to constraint in resources and time. The work has been implemented and analysed with satisfactory results.

References

[1] Roberto Cruj, Nancy Lopez, A parallelization of a heuristic for the maximum Clique problem, Journal of Computing Sciences in Colleges, Vol. 17 , pp 62-70, 2001.

[2] R.Cruz and N. Lopez, A Neural Network Approach to Solve the Maximum Clique Problem, In Proceedings of the V European Congress on Intelligent Techniques and Soft-Computing - EUFIT, Vol. 1, pp. 465-470, 1997.

[3] Coen Bron, Joep Kerbosch, Finding all cliques in an undirected graph, Communications of the ACM, Vol. 16, pp 575-577, 1973.

[4] James Cheng, Yiping Ke, Fast Algorithms for Maximal Clique Enumeration with Limited Memory, ACM 978-1-4503-1462-6, 2012.

[5] David R. Wood, An Algorithm for Finding a Maximum Clique in a Graph, Operations Research Letters, Vol. 21 , pp. 211-217

[6] Panos M. Pardalos, A exact algorithm for the Maximum clique problem, Operations Research Letters, Vol. 9, pp 375-382, 1990.

[7] Harsh Bhasin, R. Mahajan, Genetic Algorithm Based Solution to Maximum Clique Problem, International Journal of Computer Science and Engineering, Vol. 4, 2012.

[8] Harsh Bhasin, Gitanjali Ahuja, Harnessing Genetic Algorithm for Vertex Cover Problem, International Journal of Computer Science and Engineering, Vol. 4, 2012.

[9] Michael R. Garey and David S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman,1979

[10] P.M. Pardalos and J. Xue, The Maximum Clique Problem, Journal of Global Optimization, Vol. 4, pp. 301-328, 1994.

[11] Etsuji Tomita, Hiroaki Nakanishi Polynomial Time solvability of the Maximum Clique problem, In Proceeding ECC'09 Proceedings of the 3rd international conference on European computing conference, pp 203-208.

[12] F. Gavril, Algorithms for a maximum clique and a maximum independent set of a circle graph, Networks,Vol. 3, pp. 261-273, 1973.

[13] Etsuji Tomita, Akira Tanaka, Haruhisa Takahashi, The worst-case time complexity for generating all maximal cliques and computational experiment Journal Theoretical Computer Science, Computing and combinatorics, Vol. 363 , 2006.

[14] K. Makino, T. Uno, New algorithms for enumerating all maximal cliques, Lecture Notes in Computer Science, Vol. 3111, pp. 260-272, 2004.

[15] T. Nakagawa, E. Tomita, Experimental comparisons of algorithms for generating all maximal cliques, Technical Report of IPSJ, MPS-52, pp. 41-44, 2004

[16] Etsuji Tomita, Toshikatsu Kameda, An efficient Branch and Bound algorithm for finding a maximum clique with computational experiment, Journal of Global Optimization Vol. 37, pp. 95-111, 2007.

[17] Harsh Bhasin, Nishant Gupta, Randomized Algorithm Approach for Solving PCP, International Journal Of Computer Science and Engineering, Vol. 4, 2012.

[18] Harsh Bhasin, Neha Singla, Modified Genetic Algorithms Based Solution to Subset Sum Problem, International Journal of Advanced Research in Artificial Intelligence, Vol. 1, No. 1, 2012.

[19] Harsh Bhasin, Neha Singla, Genetic based Algorithm for N – Puzzle Problem, International Journal of Computer Applications, Vol. 51, No.22, 2012.

[20] Bhattacharya B.K,Kaller D, An O(m+nlogn) Algorithm for the Maximum clique problem in circular arc graphs, Journal of Algorithms, Volume 25 ,pp 336 – 358, 2007.

[21] Harsh Bhasin, Nakul Arora, Cryptography using Genetic Algorithms, In Reliability Infocom Technology and Optimization, Faridabad, 2010.

[22] Harsh Bhasin and Neha Singla, Harnessing Cellular Automata and Genetic Algorithms to Solve Travelling Salesman Problem, International Conference on Information, Computing and Telecommunications (ICICT), conference proceedings, pp 72 – 77, 2012.