Hybrid Genetic Algorithm for Maximum Clique Problem

Harsh Bhasin¹, Naresh Kumar² and Deepkiran Munjal³

¹Delhi Technological University
²M.Tech Scholar, BSAITM, Faridabad
³Assistant Professor, BSAITM, Faridabad

ABSTRACT

Maximum clique problem is one of the most important NP-hard problems which find its applications in numerous fields ranging from networking to the determination of the structure of a protein molecule. The present work carries forward one of our earlier works and examines the effect of variation of parameters for achieving optimization. The results obtained have been presented in the paper and are extremely encouraging. The premise used in the work is Genetic Algorithms, owing to their exceptional searching capabilities. The technique can be extended to many unsolvable problems and can be used in many applications from loop determination and circuit solving.

Keywords: Genetic Algorithms, NP Hard Problems, Maximum clique problem, Searching techniques.

1. INTRODUCTION

The maximum clique problem refers to finding out the maximum clique in the graph G=(V,E), where V represents the set of vertices and E=(x,y): x,y ∈ V. Finding out a clique, in a graph, is one of the original NP hard problems. It is not possible to craft a P algorithm for the problem. The problem is immensely important as it is used in error correcting codes and tilling geometry problems. It has recently been established that the matching of 3-D structure is important in the modeling of a protein structure. Since, maximum clique problem helps in recognition of cliques, therefore the problem reduction approach may help us to map the maximum clique problem to the determination of protein structure. The premise used to obtain solutions of the problem, in the work, is Genetic Algorithms. Genetic algorithms are known for their ability to find out the best possible solution from amongst a given set, if optimization is to be achieved. The present work correlates maximum clique problem with Genetic Algorithms to find an optimal solution. The work is an extension of one of our earlier works [3]. In our earlier work the crossover rate was taken as 2% and mutation rate as 0.5%. In the present work, the parameters have been varied and the earlier algorithm has been modified. The results obtained are much better than earlier results. In order to instill confidence in the proposed technique an extensive literature review has been carried out to find the gaps in the existing methodologies. The proposed technique has been implemented in C#, .net frame work 4.0 and gives encouraging results. The rest of the paper has been organized as follows. Section 2 of the paper presents the literature review. Section 3 explains the Genetic Algorithms. Section 4 presents the proposed technique. Section 4 discusses the experimental verification and finally, section 5 discusses the conclusions and future scope.

2. LITERATURE REVIEW

An extensive literature review has been carried out in order to find the gaps in the exiting technique and to propose the new technique. The literature review has been carried out in accordance with guidelines proposed by Kitchenham [23]. It may be stated that the papers of high quality journals have been selected and as far as possible the review has been unbiased. An extensive review of the existing techniques forms the foundation stone of a good research. Therefore an extra effort has been made in order to carry out the appraisal. The results of the review have been summarized in Table 1.

Table 1: Review of Maximum Clique Problem
3. **GENETIC ALGORITHM**

Genetic algorithms (Gas) are used for optimization problem. GAs is currently being used in disciplines ranging from in engineering to social science. The current work uses a version of simple genetic algorithm. The three operators are mutation, crossover and replication. The use of GAs in optimization problem has been extensively explored by Goldberg. He applied genetic algorithm in natural gas pipeline and control the flow is governed by transition equation. The flow depends on the difference in pressure. It may be noted that the crossover and mutation operator cannot be applied in an arbitrary manner. This point was also observed by Goldberg. He chose crossover rate as 1% and mutation rate as 0.1% in the problems stated above. For applying Gas, encoding is done. The problem at hand is depicted a set of chromosomes. After the encoding the two operator crossover and mutation is applied. Crossover is applied to amalgamate the feature of the existing population whereas mutation is applied to break the local maxima. Broadly GAs can be described by the following steps.

1. **Population Generation:** A population of chromosomes is generated. The population is generally binary.
2. **Crossover:** From amongst the above chromosomes, two chromosomes are selected and a crossover point is taken with the help of which the new population is generated. This is referred to as one point crossover. However two point...
crossover and multipoint crossover are also prevalent.

3 Mutation: - Mutation refers to the flipping of any random bit of random chromosomes to generate a new set of chromosomes.

4 Reproduction: - Reproduction operator helps us to replicate the fit chromosome again and again so as to increase the average fitness of the population.

4. MAXIMUM CLIQUE PROBLEM

Maximum clique problem is an NP complete problem. The problem calls for finding out all the fully connected subgraphs from given graph. The concept is not just used for mathematical problems but also in circuit theory. It may be noted that, a clique corresponds to individual set in the graph. The clique problem is related with a given group of vertices some of which have edges between them. The maximal clique is the chief subset of vertices in which each spot is unwaveringly connected to every other apogee in the subset.

5. PROPOSED WORK

Step 1: An initial population consisting of m chromosomes is taken. Each chromosome has n cells where n is the number of vertices in graph. The population is binary.

Step 2: Each chromosome consist of 0s and 1s. The chromosomes are then mapped to the vertices of the graph.

Step 3: Each chromosome is assigned an index based on the above step. The index will be more if the number of vertices selected is less and the sub graph is fully connected. We did not take into account the second condition in our previous work. Thus leads to misleading results.

Step 4: For each chromosome, fitness value is calculated according to formula 1/ (1+e - multiplied by a constant. The constant is determined by regression analysis.

Step 5: A threshold is decided which is taken half the range of indices obtained in step 3.

Step 6: Crossover and mutation are performed and new chromosomes are subjected to step 3 and 4. Here the difference from our previous work is that the crossover rate is taken as 2% and the mutation rate as 0.5% for optimal results.

Step 7: Reproduction is carried out with Roulette wheel Selection.

Step 8: The above process can be repeated and the results can be compared with previous iteration in order to improve solution. Here, the previous results are stored and act as a guided weight for the next step. The process is depicted in the following diagram.

Figure 1: Maximum Clique Problem
6. RESULT AND CONCLUSION

In the above work, it has been again established that, applying the above technique results in the crafting of a set of solutions that provides almost certain answer to the problem. The Vertex Cover is an NP Hard problem. The method proposed has been explained in one of our earlier works and is verified again by varying the parameters of Genetic Algorithms. The algorithm has now been tested sufficient number of times. The above method has also undergone regress testing. We can now safely state that the above novel approach, that uses the blend of nature along with technical novelty, is capable of producing requisite results. It may be stated here that, this technique is being used to solve other NP hard problems as well.

References