

# SQL Query Optimization Techniques

**Bhagyashri S. Bharambe<sup>1</sup>, Priyanka R. Munot<sup>2</sup>, Dipali R. Patil<sup>3</sup>, Kajal P. Pathak<sup>4</sup>**

Department of Computer Engineering, SSBT Collage of Engineering, Jalgoan, India<sup>1,2,3,4</sup>

**Abstract:** Query optimization is the overall process of choosing the most efficient means of executing a SQL statement. The optimizer attempts to generate the best execution plan for a SQL statement. The best execution plan is defined as the plan with the lowest cost among all considered candidate plans. SQL is a nonprocedural language, so the optimizer is free to merge, reorganize, and process in any order. The cost is a number that represents the estimated resource usage for an execution plan. The cost computation accounts for factors of query execution such as I/O, CPU, and communication. To implement query optimization methods such as Heuristic Greedy based optimization, Iterative Improvement based cost optimization and Ant Colony optimization algorithms. Show Comparison of cost, execution time and response time between Heuristic Greedy based optimization; Ant Colony Optimization and Iterative Improvement based cost optimization algorithms

**Keywords:** Optimization, Genetic, Ant colony algorithm, Heuristic

## I. INTRODUCTION

The query optimization is a process of finding the most effective execution plan for the given user submitted query. Query optimization has been found very useful in increasing the database systems' performance in terms of time. A query written in a high level language needs to convert into a form that system can understand and perform further processing. In its internal form, the relational algebra expression, there are number of variations available for representation. Also the various query optimization strategies and algorithms are available to compute the answer. Researchers have worked with various operations of the query to find out the most efficient query execution plan, various techniques to choose the optimal solution among the various methods, etc. Different query optimization techniques have been applied like rule based optimization, cost-based query optimization, deterministic optimization, randomized optimization and their variations. Results of query optimization can be used by different emerging database management systems. The database users can get benefits of the optimization by getting the results to the query in a timely and predictable manner. Database vendors can use them to improve the efficiency of their DBMS which will provide support to the upcoming huge amount of data.

## II. LITERATURE SURVEY

'**N. Satyanarayana** ': proposed a new dynamic query optimization algorithm based on the greedy algorithm that uses the randomized strategies. The execution cost of queries and system resources requirements were reduced significantly and applicable to both distributed and centralized database systems.

'**Pravin Chandra, Anurag Jain** ': The general query optimization techniques like CBO, RBO. Also presented the techniques used by the Oracle. Surajit Chaudhuri and Kyuseok Shim [11] proposed greedy conservative heuristic as a technique to optimize single block of SQL with group-by. The implementation was with a System-R style optimizer. This approach extended the traditional optimization algorithms for Multi-block queries using pull-up as well as pull-down transformations. The paper also discussed the join-aggregate class of nested queries and queries containing views with aggregates.

'**Michael C. Ferris and Yannis E**': University of Wisconsin and obtained significant computational savings over the randomized methods by parallel implementation. A set of different queries, with size of each query consisting up to 16 joins, was tested on System-R algorithm and GA. The experiment found that GA works relatively better than System-R optimizer.

## III. ALGORITHMS

**A. Genetic Algorithm:** The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that is based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next



generation. Over successive generations, the population "evolves" toward an optimal solution. You can apply the genetic algorithm to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, non differentiable, stochastic, or highly nonlinear. The genetic algorithm can address problems of mixed integer programming, where some components are restricted to be integer-valued.

**B. Heuristic Algorithm** A heuristic algorithm is one that is designed to solve a problem in a faster and more efficient fashion than traditional methods by sacrificing optimality, accuracy, precision, or completeness for speed. Heuristic algorithms are often used to solve NP-complete problems, a class of decision problems. In these problems, there is no known efficient way to find a solution quickly and accurately although solutions can be verified when given. Heuristics can produce a solution individually or be used to provide a good baseline and are supplemented with optimization algorithms. Heuristic algorithms are most often employed when approximate solutions are sufficient and exact solutions are necessarily computationally expensive.

**C. Ant Colony Algorithm:** In computer science & operations research, the Ant Colony Optimization Algorithm (ACO) is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. Artificial Ants stand for multi-agent methods inspired by the behavior of real ants. The pheromone-based communication of biological ants is often the predominant paradigm used.<sup>[2]</sup> Combinations of Artificial Ants and local search algorithms have become a method of choice for numerous optimization tasks involving some sort of graph, e.g., vehicle routing and internet routing. The burgeoning activity in this field has led to conferences dedicated solely to Artificial Ants, and to numerous commercial applications by specialized companies such as AntOptima.

#### IV. PROPOSED SYSTEM

Query optimization is the part of the query process in which the database system compares different query strategies and chooses the one with the least expected cost. The query optimizer, which carries out this function, is a key part of the relational database and determines the most efficient way to access data. It makes it possible for the user to request the data without specifying how these data should be retrieved. The cost of accessing a query is a weighted combination of the I/O and processing costs. The I/O cost is the cost of accessing index and data pages from disk. Processing cost is estimated by assigning an instruction count to each step in computing the result of the query

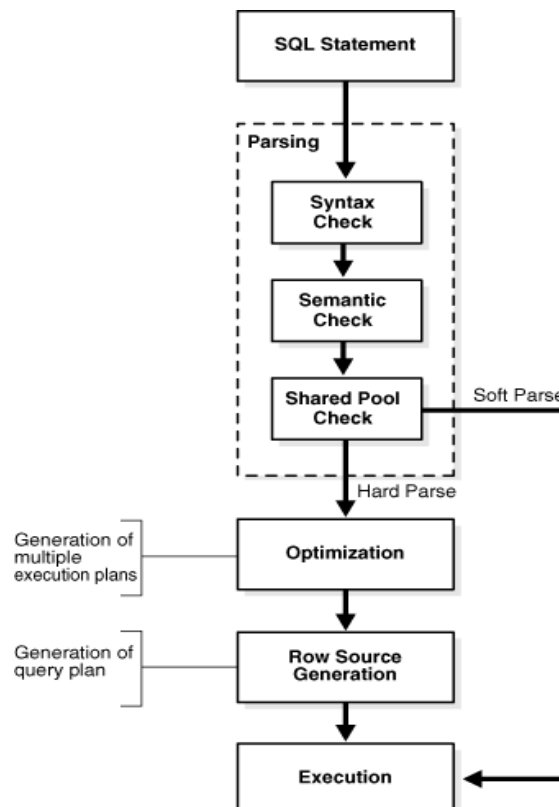


Figure 2. SQL query processing

V. RESULT

This is the homepage, in that the name of tables are included and the columns are also displayed. We want to give some SQL query and then the result in different ways:

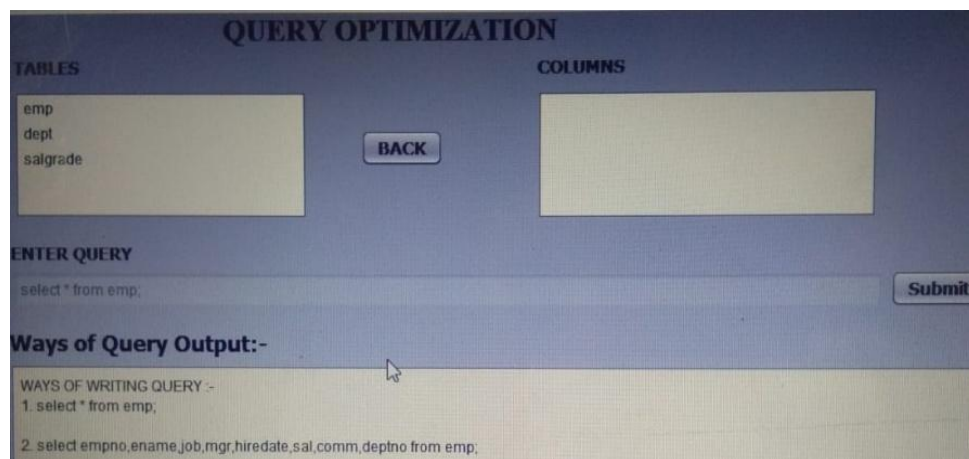
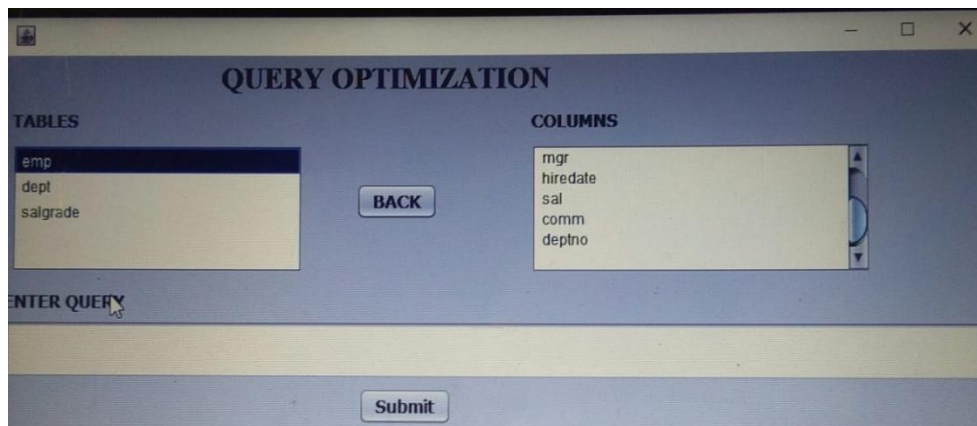


Figure 5. Ways of query

empno	ename	job	mgr	hiredate	sal	comm	deptno
7839	KING	PRESIDENT		1981-11-17	5000		10
7698	BLAKE	MANAGER	7839	1981-05-01	2850		30
7782	CLARK	MANAGER	7839	1981-06-09	2450		10
7566	JONES	MANAGER	7839	1981-04-02	2975		20
7788	SCOTT	ANALYST	7566		3000		20
7902	FORD	ANALYST	7566	1981-12-03	3000		20
7369	SMITH	CLERK	7902	1980-12-17	800		20
7499	ALLEN	SALESMAN	7698	1981-02-20	1600	300	30
7521	WARD	SALESMAN	7698	1981-02-22	1250	500	30
7654	MARTIN	SALESMAN	7698	1981-09-28	1250	1400	30
7844	TURNER	SALESMAN	7698	1981-09-08	1500	0	30
7876	ADAMS	CLERK	7788		1100		20
7900	JAMES	CLERK	7698	1981-12-03	950		30
7934	MILLER	CLERK	7782	1982-01-23	1300		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10
7839	KING	PRESIDENT		1981-11-17	5000		10

Figure 6. Table of output

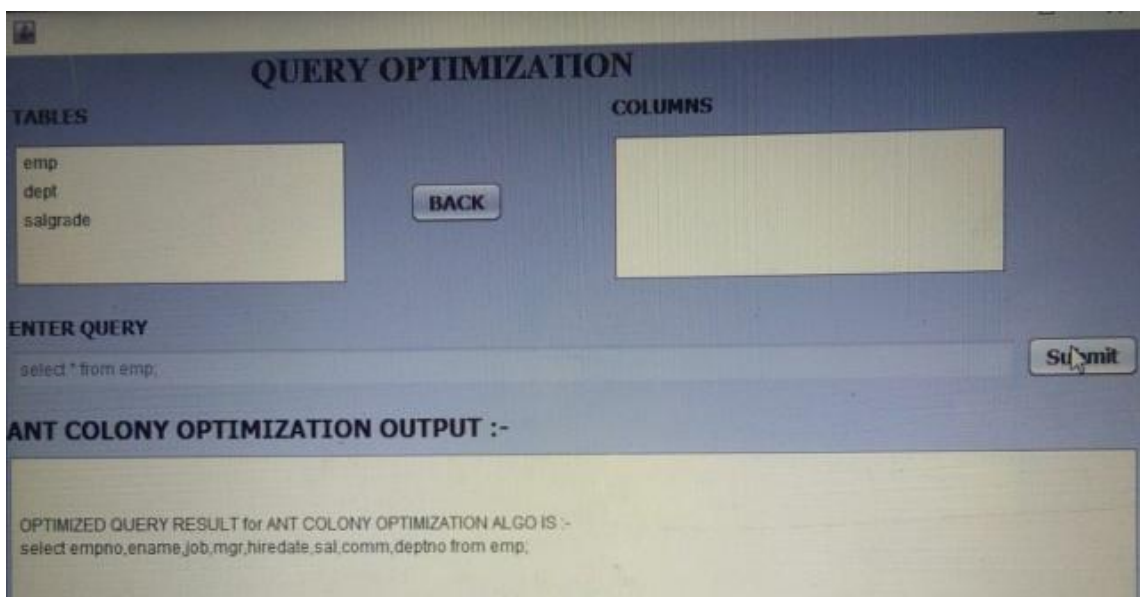


Figure 7: Output of Ant Colony Algorithm

## VI. CONCLUSION

In this, we got the Optimized query by using algorithm like Genetic, Heuristic and Ant colony ..We have also calculated the CPU load and Execution time.

## REFERENCES

- [1]. N. Satyanarayana, SK.Sharfuddin, SK.Jan Bhasha, "New Dynamic Query Optimization Technique In Relational Database Management Systems", International Journal of Communication Network Security, ISSN: 2231 – 1882, Volume- 2, Issue-2, 2013
- [2]. Pravin Chandra, Anurag Jain, and Manoj Kr. Gupta, Query Optimization in Oracle, "Voyager-The Journal of Computer Science and Information Technology", ISSN 0973- 4872, Vol. 6, No.1, p.p. 18-22, July-Dec. 2007
- [3]. Michael L. Rupley, Jr., "Introduction to Query Processing and Optimization", 2004
- [4]. Majid Khan and M. N. A. Khan, Exploring "Query Optimization Techniques in Relational Databases, International Journal of Database Theory and Application" Vol. 6, No. 3, June, 2013
- [5]. Michael Stillger and Myra Spiliopoulou, "Genetic Programming in Database Query Optimization".
- [6]. Dr. P.K.Butey, Prof. Shweta Meshram and Dr. R.L. Sonolikar, "Query Optimization by Genetic Algorithm", Journal of Information Technology And Engineering Vol.3 No.1 Jan-June 2012 pp. 44-51 ISSN: 2229-7421