# Enhanced Multi-Objective Ant Colony Based Query Optimization Technique

Daljinder Dugg[1], Er. Mandeep Singh[2], Dr. Gurpreet Singh[3]
M.Tech Scholar[1], Assistant Professor[2], Professor[3]
Department of Computer Science & Engineering
St. Soldier Institute of Engineering &Technology, Jalandhar, India

**Abstract:**
Query optimization is a stimulating task of any database system. The results of Entropy Based Restricted Stochastic Query Optimizer (ERSQO) are compared with the results of Exhaustive Enumeration Query Optimizer (EAQO), Simple Genetic Query Optimizer (SGQO), Novel Genetic Query Optimizer (NGQO) and Restricted Stochastic Query Optimizer (RSQO). In terms of Total Costs, EAQO outperforms SGQO, NGQO, RSQO and ERSQO. However, stochastic approaches dominate in terms of runtime.To overcome the issues associated with the existing techniques, a new multi-objective ant colony based query optimization technique is proposed. The effect of query cost and communication overheads will also be considered.The use of ant colony optimization can find optimistic query in order to reduce the query cost.

**Keywords:**Distributed database**,** Query optimization, Ant colony based otimization.

## 1. INTRODUCTİON

Distributed database techniques supply a noticeable difference with communication along with data producing due to its data circulation all over n/w sites. But not only is actually data gain access to faster but a new single-point connected with inability is actually more unlikely that that occurs along with it offers local regulate of data regarding users. Nevertheless, at this time there is some intricacy any time trying to handle along with regulate handed out collection systems. Spread circle computing conditions have grown to be a new cost-effective along with well-liked preference to accomplish high performance along with in order to resolve large computational problems. As opposed to previous supercomputers, a new handed out collection computing program can be used multi-purpose computing program to run diversified high performance similar applications. The trends around computer system web 2 . technologies along with collection techniques technologies triggered the creation of handed out directories while in the middle 1970s. It was believed that many apps is handed out sometime soon and for that reason, the particular directories must be handed out also. Any handed out collection program features a Spread Data source Management Program (DDBMS), a new handed out collection including a circle regarding interconnections. The objective of a new DDBMS will be to regulate the particular therapy for a new Spread Details Platform (DDB) such this seems like in order to the person to be a dierected database.

### 1.1 Components of Distributed Database
DDBMS comprises of following components:
**1. Database Manager**: is the software responsible for processing a segment of the distributed database
**2. Distributed Database Management System:** is defined as the software which governs a Distributed Database System. It supplies the user with the illusion of using a centralized database.
**3. User Request Interface:** known some times as a customer user interface, which is usually a client program that acts as an interface to the distributed transaction manager. A customizable

user interface is provided for entering requested parameters related to a database query. The customized parameter user interface provides parameter entry dialogs/windows in correlation to a data view (e.g. form or report) that is produced according to a database query.

**4. Distributed Transaction Manager:** is actually software in which translates asks from the person within actionable asks for the collection supervisors which can be ordinarily distributed. A distributed database method consists of their Distributed Transaction Manager (DTM) as well as the Data Base Manager (DBM).

### 1.2 Distributed Query Processing
Query processing deals with creating algorithms that review concerns in addition to transform these people into several files mind games operations. The issue is how you can go with a technique for doing just about every issue above the multi-level throughout the most cost effective way. The factors to be considered are classified as the distribution of info, interaction expenditures in addition to absence of sufficient in area offered information.

### 1.3 Query Optimizatıon
Query optimization is a purpose of lots of relational collection control systems. The query optimizer endeavours to view the best means to execute a certain problem by looking at the potential problem plans. Usually, a problem optimizer are not seen directly by customers: the moment requests are sent to collection machine, and also parsed because of the parser, they may be after that approved to your problem optimizer exactly where optimization occurs.

### 1.4 Multi-Objective Query Optımizatıon
1. There are many cases are some other value analytics besides performance time which can be relevant to assess dilemma plans.
2. Throughout cloud computing, by way of example, you will need to assess query plans not just with regard to the time people

consider to try and do and also with regard to what quantity of money their own performance costs.

3. The context regarding approximate dilemma search engine marketing, it is easy to do dilemma plans on randomly decided on samples in the enter details to be able to receive approximate final results by using lower performance overhead.

4. When this happens, alternative dilemma plans has to be when compared to with regard to their own performance time and also due to the detail or longevity of your data people generate.

## 1.5 Ant Colony Optimization

In ACO, a number of artificial ants help in building solution for optimization problems and via a communication scheme, they exchange the information. They find the shortest paths as the moving ants lay pheromone on the ground, so that when another ant encounters it, it can detect it and decide to follow the trail. As a result, the emerged collective behavior is an indication that if a number of ants choose a particular path, then the probability of other ants following the same path increases. The main idea of ACO is to model a problem as the search of minimum cost path in a   graph.

---

1. Begin
2. Initialize pheromone trails;
3. Produce an underlying population of solutions (ants);
4. For every ant $f \epsilon a$: calculate fitness function (f);
5. Determine the best position for each ant;
6. Determine the best global ant;
7. Update the pheromone;
8. Determine if termination= true;

---

## 2. RELATED WORK

**Manik Sharma avec al. (2016) [1]** query marketing is actually a stimulative endeavor associated with a database system. Quite a few heuristics have been applied to recent years, which recommended brand-new algorithms intended for substantially helping the performance of the query. The particular look for a better solution even now continues. Right here, a creative stochastic framework involving DSS issue optimizer can be recommended to help promote boost the design of pre-existing issue marketing genetic approaches. **Zhan Li avec al. (2016) [2]** properly designed data structure known as Reference point Key Major dining room table (RPK-table) which stores the relationship involving major key as well as dangerous key among tables. According to this particular structure, one can propose to your girlfriend a better formula with Guide Decrease framework intended for join-aggregate query. Experiments with TPC-H dataset exhibit that our formula outperforms pre-existing solutions in terms of communication price as well as issue effect time.**Varghese S. Chooralil avec al. (2015) [3]** outlined the requirement for a few retracts structure to help method a category involving smartly designed end user issues based on Reuters-21578 Textual content Categorization Variety Facts Established, known as Semantic Reference Information having Compound Method (SRD-CP). The primary fold constructs a new Semantic Pattern Sapling using the Reference Information Structure (RDF) issue words intended for uncovering the papers subject matter name. The particular RDF issue words makes use of mixture structure to produce the results to the widened end user queries. Your second fold can be the design of mixture structure around SRD-CP framework this contains HTTP project as well as constrained use project to handle the difficult issue processing. Lastly, the 3rd fold features the phrase concept co-occurrences utilizing the

connection volume variety by making use of the constructed SPT.**Fuqi Tune avec .(2014) [4]** have got recommended a strategy intended for doing issue setting up as well as marketing based on a prolonged issue design chart as well as heuristics. 1st, this particular newspaper generalizes SPARQL issue statement manifestation if you take various other expression note, aiming on alleviating the constraints of only employing primary issue three times the patterns. Second, this particular newspaper is the heuristics intended for price the price tag on making issue three times the pattern. The particular recommended issue setting up solutions usually are carried out within Corese issue motor and they are analyzed employing BSBM benchmark. The outcomes report that the recommended solutions could boost successfully the issue delivery period of SPARQL issue engine.**Chen Yan avec al. (2012) [5]** distributed database system, shown the desired goals involving distributed database issue marketing, as well as examined the issue marketing method based on semi-join business with the simple application. Furthermore, them presented a new established formula used intended for multiple network as well as issue marketing using the semi-join issue marketing, the SDD-1 algorithm. **Philip Henry Beran avec al. (2011) [6]** Having the arrival connected with allocated computing, notably ever since the introduction connected with Grids, Atmosphere and various Program Oriented Computing paradigms, the particular querying of huge datasets connected with allocated data source or maybe details repositories for a global range in to a challenging analysis question. At present, beside various themes, not one but two significant worries in this particular analysis region must be resolved: details gain access to & plug-in and query delivery planning. Each of our analysis effort details the second situation, that is the particular query optimization connected with allocated data bank queries. Hereby we all take into account a range of diverse heterogeneous and homogeneous infrastructures, similar algorithms, and large datasets, which will course across numerous exclusive businesses (VOs) by using usually simply no centralized authority. This kind of newspaper presents a new new heuristic platform to the optimization connected with query delivery ideas (QEP) using a world-wide scale. Each of our efforts are based on a multi-staged blackboard apparatus to determine which obtainable details, options and operations must be viewed as to execute a query optimally. Furthermore, the evaluate predicament attests our studies that perhaps little modifications in the selection of e.g. type businesses intended for a question delivery woods (QET) bring on major effectiveness improvements.**Zhang avec al. (2011) [7]** made a cost optimization type intended for multi-role-based determination service system (DSS). Based on investigating this company pass and dealing technique of ore working, we all offer the buildings connected with DSS that's made dependant on multi-roles. This kind of DSS structure pre-processes the results intended for products and aspects, generates an overall data bank, abstracts the related optimum businesses analysis styles and presents the particular reasoning apparatus connected with a professional system. A non-linear kind of ore working intended for blast heaters and it is solutions usually are provided.**Abu-Naser avec al.(2011) [8]** currently have submitted DSS intended for offering the understanding intended for optimizing the particular lately followed e-learning education strategy throughout helpful institutions.The suggested DSS draws on discovery (mining) of information from a lot of info produced in the doing work the particular organization to the business. This knowledge enables you to information and enhance any kind of start up business strategy put in place from the institution. A suggested DSS entails

Databases motor, Information Mining motor and Artificial Cleverness engine. Most of these motors band together so as to acquire the skills important to boost the potency of any kind of strategy, as well as e-learning.

## 2.1 Gaps in literature

As discussed by Manik Sharma [1] , proposed Query optimization is really a revitalizing process of any database system. Many heuristics are actually applied to modern times, that recommended brand new algorithms for extensively increasing the overall performance of any query. The actual try to find the answer nonetheless continues. By simply conducting your review, that is located that will the existing researchers have overlooked several issues.The effect of query cost and communication overheads are ignored in most of existing research on distributed databases as well as use of multi-objective optimization is ignored by most of existing researchers and use of ant colony optimization to reduce query cost is also neglected in existing literature.

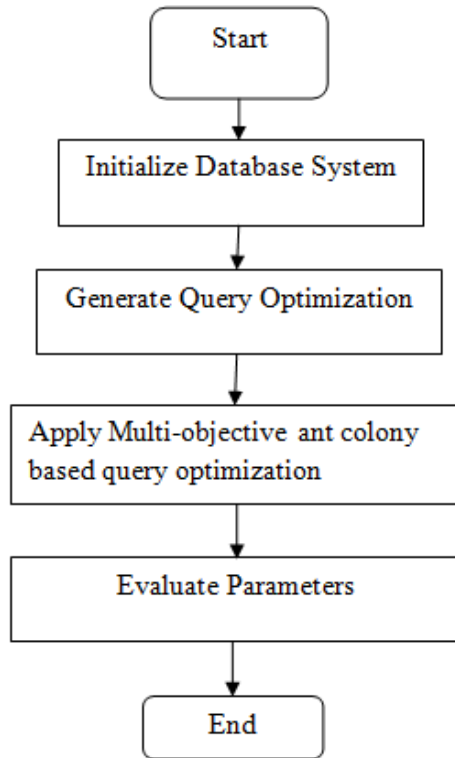## 3.METHODOLOGY

### 3.1 Proposed methodology



Figure.2.proposed methodology

## 3. 2 Experimentation And Results

In this first objective is achieved i.e. to evaluate the performance of existing query optimization based techniques. Matlab is high performance language for technical computing and weka is a collection of machine learning algorithms for data mining task.

1.      **Query Cost:** Cost    is generally calculated when whole past time period with regard to giving answers to query.
2.      **Communication        Overheads** Communication Overhead could be the percentage of time you spend getting in touch with ones crew preferably of asking for fruitful work done.
3.      **Response Time:** Response time is the overall period of time it will require to respond to the ask services that will service is usually everything originating from a memory get, to a disk IO, to intricate repository question or perhaps filling the whole webpage.
4.      **Execution Time:** Execution time is the time using a person's software is exercising and also executing. As opposed to diverse program life-cycle steps as well as collect time frame, website page web page link considerable degrees of heap time.
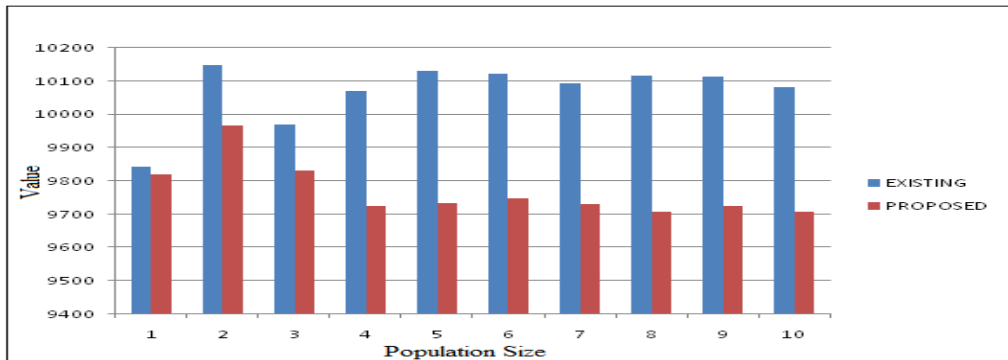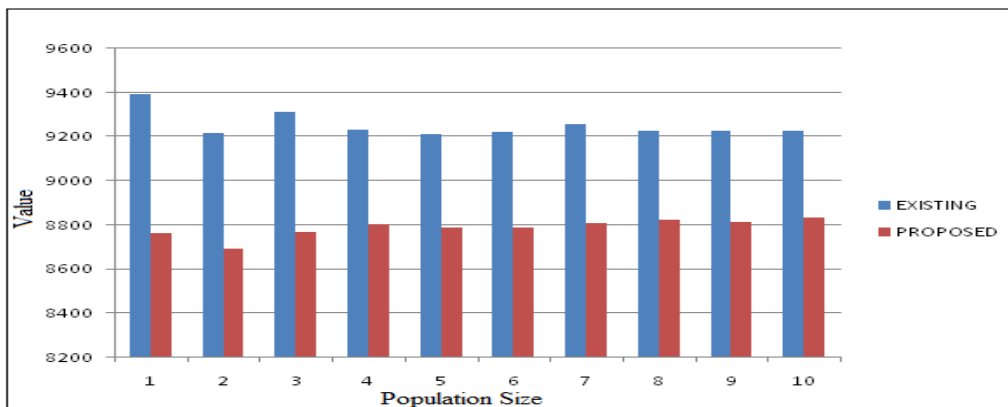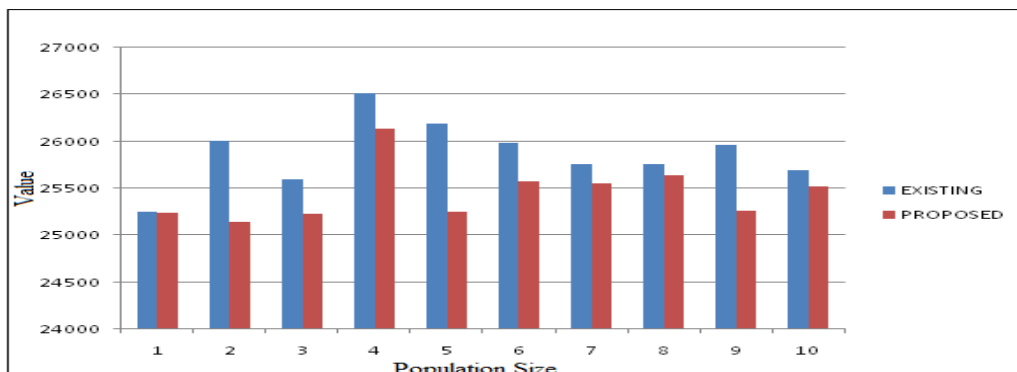
### 1. Over head
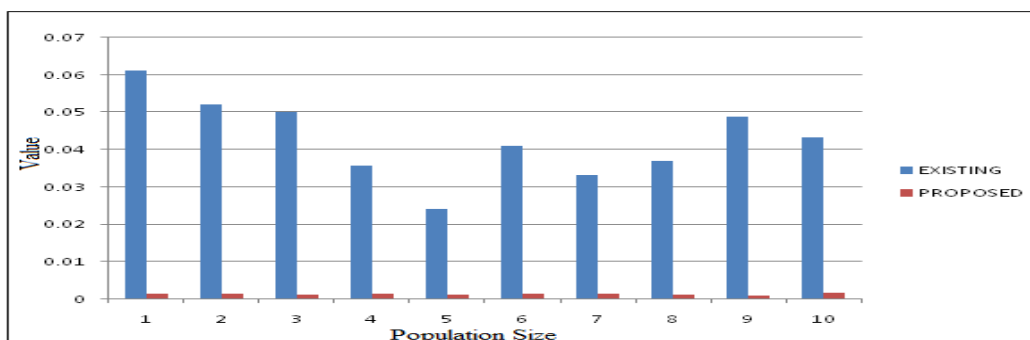
**2. Execution**
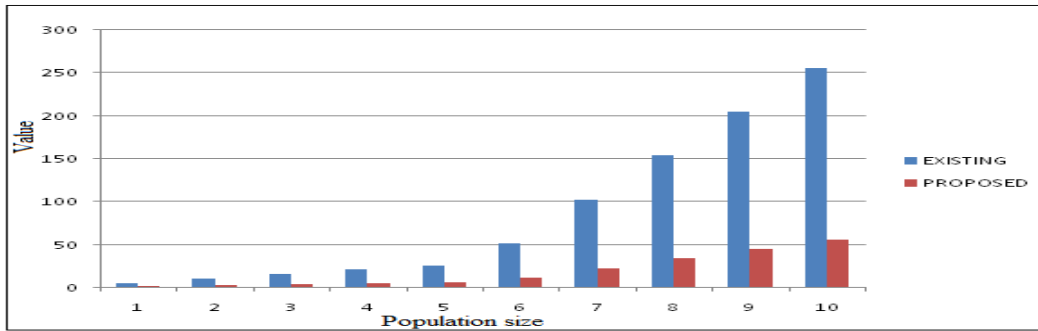


**3. IO Cost**



**4. CPU cost**



**5.    Total Cost**



## 4. CONCLUSİON AND FUTURE WORK

This paper represnts that Query optimization is actually a purpose of several relational data bank supervision systems. The outcomes associated with Entropy Based Restricted Stochastic Issue Optimizer (ERSQO) tend to be compared with the outcome associated with Exhaustive Enumeration Issue Optimizer (EAQO), Uncomplicated Inherited Issue Optimizer

(SGQO), Book Inherited Issue Optimizer (NGQO) along with Restricted Stochastic Issue Optimizer (RSQO). With regard to Whole Charges, EAQO outperforms SGQO, NGQO, RSQO along with ERSQO. Nonetheless, stochastic approaches dominate regarding runtime. The actual Whole Charges manufactured by ERSQO is preferable to SGQO, NGQO along with RGQO by simply 12%, 8% along with 5% respectively. To conquer the down sides linked to the existing procedures, a new

multi-objective insect nest dependent query marketing technique is proposed. The issue associated with query charge along with transmission overheads can also be considered.The use of insect nest marketing can find beneficial query so as to reduce the query charge along with the consist of method might be compared with existing procedures considering specified overall performance metrics.but just after there are a few problems that is improved with forseeable future with the far more evolutionary marketing technique.

## 5. REFRENCES:

[1].Sharma M., Singh M. and Singh R., Design and analysis of stochastic DSS query optimizers in a distributed database system, Egyptian Informatics Journal. 17(2016) 161-173.

[2].Li Z., Feng Q., Chen W. and Wang T., RPK-table based efficient algorithm for join-aggregate query on MapReduce, CAAI Transactions on Intelligence Technology. 1(2016) 79-89.

[3].Chooralil V.S. and Gopinathan., A Semantic Web query Optimization Using Resource Description Framework, Procedia Computer Science. 70(2015) 723-732.

[4].Song F. and Corby O., Extended Query Pattern Graph and Heuristics - based SPARQL Query Planning, Procedia Computer Science. 60(2015) 302-311.

[5].Yan C., Lin Z., Taoying L. and Yinging Y., The semi-join query optimization in distributed database system. International Conference on Information Technology and Computer Science. Atlantis Press. 2012, pp. 606–609.

[6].Beran P.P., Mach W., Schikuta E. and Vigne R., A Multi-Staged Blackboard Query Optimization Framework for World-Spanning Distributed Database Resources, Procedia Computer Science. 2011, pp. 156-165.

[7].Gowda K.A., Nath M. and Jayaram., Application of genetic algorithm optimized neural network connection weights for medical diagnosis of Pima Indians diabetes. International Journal on Soft Computing. 2(2011) 15–23.

[8].Zhang, Ruijun, Lu J. and Zhang G., A knowledge-based multi-role decision support system for ore blending cost optimization of blast furnaces. European Journal of Operational Research. 1(2011) 194-203.

[9].Naser A., Masri A. A., Sultan Y. A. and Zaqout I., A prototype decision support system for optimizing the effectiveness of e-learning in educational institutions. International Journal of Data Mining & Knowledge Management Process.  4 (2011) 1-13.

[10].Zhengmeng, Chai, and Haoxiang J., A brief review on Decision Support Systems and it's applications. In IT in Medicine and Education (ITME), International Symposium. 2(2011) 00401-405.

[11].Chaudhuri, Surajit, Giakoumakis L., Narasayya V. and Ramamurthy R., Rule profiling for query optimizers and their implications, In Proc of IEEE 26th International Conference on Data Engineering (ICDE).  2010, pp. 1072-1080.

[12].Pyar and Kyi., Decision support system for personnel information using data warehouse, International Conference on Computer and Automation Engineering (ICCAE). 1(2010) 668-672.

[13].Hagmann, Robert, and Skeen M. D., Real-time query optimization in a decision support system. U.S. Patent 6, 338, 055  issued. 2002.

[14].Bara, Adela, Diaconita V., Lungu I. and Velicanu M., Improving performance in integrated DSS with object oriented modelling, WSEAS Transactions on Computers. 4 (2009) 599-609.

[15].Andersch, Christian, Coates M. L., and Saueressig G., Methods to browse database query information. U.S. Patent 7,447,687 issued. 2008.

[16].Mansmann, Svetlana, and Scholl M. H., Decision support system for managing educational capacity utilization, IEEE Transactions on Education 50. 2 (2007) 143-150.

[17].Damart, Sébastien, Dias L. C., and Mousseau V., Supporting groups in sorting decisions: Methodology and use of a multi-criteria aggregation/disaggregation DSS, Decision Support Systems 43. 4(2007) 1464-1475.

[18].Zhang, Shifeng, and Goddard S., A software architecture and framework for Web-based distributed Decision Support Systems. Decision Support Systems.  4(2007) 1133-1150.

[19].Keenan, Peter B., Spatial decision support systems: a coming of age, Control and Cybernetics 35. 1 (2006) 9.

[20].Xiong, M. H., Tor S. B., Bhatnagar R., Khoo L. P. and Venkat S., A DSS approach to managing customer enquiries for SMEs at the customer enquiry stage, International Journal of Production Economics. 1 (2006) 332-346.