Online ISSN : 0975-4172
Print ISSN : 0975-4350

Global Journals

of Computer Science and Technology: C

Software & Data Engineering

Technique for Real-Time

Algorithm for Frequent

Highlights

Answering Top-K Queries

An Efficient Concurrency

Discovering Thoughts, Inventing Future

VOLUME 13 ISSUE 2 VERSION 1.0

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Volume 13 Issue 2 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

(A Delaware USA Incorporation with "Good Standing"; Reg. Number: 0423089)

Sponsors: Global Association of Research

Open Scientific Standards

Publisher's Headquarters office

Global Journals Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141 United States

USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Association of Research, Marsh Road, Rainham, Essex, London RM13 8EU
United Kingdom.

Packaging & Continental Dispatching

Global Journals, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investers@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:

22 USD (B/W) & 50 USD (Color)

Yearly Subscription (Personal & Institutional):

200 USD (B/W) & 250 USD (Color)

Dr. Bart Lambrecht
Director of Research in Accounting and FinanceProfessor of Finance
Lancaster University Management School
BA (Antwerp); MPhil, MA, PhD (Cambridge)

Dr. Carlos García Pont
Associate Professor of Marketing
IESE Business School, University of Navarra
Doctor of Philosophy (Management), Massachusetts Institute of Technology (MIT)
Master in Business Administration, IESE, University of Navarra
Degree in Industrial Engineering, Universitat Politècnica de Catalunya

Dr. Fotini Labropulu
Mathematics - Luther College
University of ReginaPh.D., M.Sc. in Mathematics
B.A. (Honors) in Mathematics
University of Windso

Dr. Lynn Lim
Reader in Business and Marketing
Roehampton University, London
BCom, PGDip, MBA (Distinction), PhD, FHEA

Dr. Mihaly Mezei
ASSOCIATE PROFESSOR
Department of Structural and Chemical Biology, Mount Sinai School of Medical Center
Ph.D., Etvs Lornd University
Postdoctoral Training,
New York University

Dr. Söhnke M. Bartram
Department of Accounting and FinanceLancaster University Management SchoolPh.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

Dr. Miguel Angel Ariño
Professor of Decision Sciences
IESE Business School
Barcelona, Spain (Universidad de Navarra)
CEIBS (China Europe International Business School).
Beijing, Shanghai and Shenzhen
Ph.D. in Mathematics
University of Barcelona
BA in Mathematics (Licenciatura)
University of Barcelona

Philip G. Moscoso
Technology and Operations Management
IESE Business School, University of Navarra
Ph.D in Industrial Engineering and Management, ETH Zurich
M.Sc. in Chemical Engineering, ETH Zurich

Dr. Sanjay Dixit, M.D.
Director, EP Laboratories, Philadelphia VA Medical Center
Cardiovascular Medicine - Cardiac Arrhythmia
Univ of Penn School of Medicine

Dr. Han-Xiang Deng
MD., Ph.D
Associate Professor and Research Department Division of Neuromuscular Medicine
Davee Department of Neurology and Clinical NeuroscienceNorthwestern University
Feinberg School of Medicine

Contents of the Volume

Global Journal of Computer Science and Technology
Software & Data Engineering
Volume 13 Issue 2 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Comprehensive Concurrency Control Technique for Real-Time Database System

By Md. Anisur Rahman & Md. Sahadat Hossain
Dhaka University of Engineering & Technology Gazipur, Bangladesh

Abstract - Real-time database must maintain the Temporal Consistency of the data which cannot be achieved with the conventional concurrency control techniques as they focus only on the consistency of the data. Different protocols exhibit good performance on different situations. But a single technique is inadequate to meet the demand of real-time database. To improve the concurrency control technique for real-time transactions, this paper will present a comprehensive technique that coordinates multi-version, OCC Sacrifice, Speculative Concurrency Control and 2PL-HP protocols. The presented technique uses best suited protocol based on the contention of transactions. Thus it can significantly improve the concurrency of transactions as well as increase the number of transactions.

Keywords : multi-version, speculative concurrency control, real-time transaction, temporal consistency.

GJCST-C Classification : H.2.8

A COMPREHENSIVE CONCURRENCY CONTROL TECHNIQUE FOR REAL-TIME DATABASE SYSTEM

Strictly as per the compliance and regulations of:

A Comprehensive Concurrency Control Technique for Real-Time Database System

Md. Anisur Rahman [α] & Md. Sahadat Hossain [σ]

Abstract - Real-time database must maintain the Temporal Consistency of the data which cannot be achieved with the conventional concurrency control techniques as they focus only on the consistency of the data. Different protocols exhibit good performance on different situations. But a single technique is inadequate to meet the demand of real-time database. To improve the concurrency control technique for real-time transactions, this paper will present a comprehensive technique that coordinates multi-version, OCC Sacrifice, Speculative Concurrency Control and 2PL-HP protocols. The presented technique uses best suited protocol based on the contention of transactions. Thus it can significantly improve the concurrency of transactions as well as increase the number of transactions.

Keywords : multi-version, speculative concurrency control, real-time transaction, temporal consistency.

I. Introduction

Real-time database systems are identified as having timing constraints and can be found in applications such as defence systems, Internet and multimedia applications, industrial automation, programmed stock trading and air traffic control etc.

The timing constraints of real-time database are typically specified in the form of deadlines that require a transaction to be completed by a specified time. Failure to meet a deadline can cause the results to lose their value, and in some cases a result produced too late may be useless or even harmful. So unlike traditional database Real-time databases (RTDBMS) must maintain Temporal Consistency of data. Temporal Consistency requires two main requirements: Absolute validity and relative consistency. Absolute validity is the notion of consistency between environment and its reflection in the database. Relative consistency is the notion of consistency of the data that are used to derive new data. The correctness of the system depends not only on the logical results but also on the time used to produce these results, as the transactions for their concurrent implementation has own timing constrains and dependence. That is Real-time systems are to ensure completion of more transactions within the deadline.

The conventional pessimistic concurrency control mechanism based on locking e.g. two phase locking with higher priority (2PL-HP) can assure the transactions serializability [1], so as to strongly assure the consistency of data. However, because of a high rate of restart of transactions, it cannot satisfy the need of the real-time database systems very well. The optimistic concurrency control techniques assume that the probability of any two concurrent transactions requesting the same data is not often. So it allows all operations to be performed directly whenever transactions request. But these transactions must pass through the validation checking before they are allowed to be committed to database.

The Multi-version Time-stamp Ordering (MVTO) technique is one type of optimistic concurrency control (OCC)mechanism which provides a large degree of concurrency for the transactions by maintaining multiple versions of data items [1]. So it is more appropriate for real-time database systems where the transaction has a low rate of restart and delay of cut-off time but a high degree of concurrency. It ensures transactions serializability using Time-stamp Ordering mechanism.

In OCC Broadcast Commit (OCC-BC) protocol, which is another OCC method, when a transaction commits, it notifies its intentions to all other currently running transactions [5]. Each of these running transactions carries out a check to test whether it has any conflicts with the committing transaction. If any conflicts are detected, the transaction carrying out the check immediately aborts itself and restarts. Note that there is no need for a committing transaction to check for conflicts with already committed transactions, because if it were in conflicts with any of the committed transactions, it would have already been aborted. Thus, in OCC-BC once a transaction reaches its validation phase, it ensures its commitment. Compared to OCC Forward protocol, it encounters earlier restarts and less wasted computations. Therefore this protocol should perform better than the OCC-forward protocol in meeting task deadlines. However, a problem with this protocol is that it does not consider the priorities of transactions. On the other hand, it may be possible to achieve better performance by explicitly considering the priorities of the transactions.

OCC Sacrifice (OCC -S) is another type of Optimistic method that considers the priority of a transaction in the validation check phase to determine which transaction(s) should be restarted [5]. Transaction with higher priority commits and

Author α : Department of CSE, Dhaka University of Engineering & Technology Gazipur, Bangladesh.
E-mails : anisur.rahman.duet@gmail.com, sahadat39@yahoo.com

causes to restart conflicting transactions. This method reduces wasted computation providing early restart of conflicting transactions.

In Speculative Concurrency Control (SCC) technique, conflicts are checked at every read and write operation. Whenever conflicts are detected a new version of each of the conflicting transactions is initiated. The primary version executes as any transaction would execute under an OCC protocol, ignoring the conflicts that develop. Meanwhile the new version executes as any transaction would do under a pessimistic protocol-subjected to locking and restarts. Improving the concurrency control protocol, this paper will present a new concurrency control method which adopts multi-version, OCC Sacrifice, SCC or 2PL-HP depending on the contention of transaction in the system. In this way, it can effectively improve the concurrency of transactions and increase the amount of the transactions completed within the deadline. The feasible analysis denotes that this new method is better than the traditional one on performance.

II. The Description of the Proposed Technique

Several concurrency control techniques exhibits better performance in different idiographic situations. In some cases locking protocols shows better performance but fails to meet the requirement in some other cases. So we have classified the type of transactions as well as consider the actual condition of the contention based on the time required to executer that transaction. Transactions in the Real-time database can be split into three categories according to Multi-version Time-stamp Ordering concurrency control method depending on the type of operations they performs [1]. They are:

a) Read-Only Transactions (Txn-R)

This type of transaction always read the data elements that is the maximum Timestamp with a less than or equal Txn-R in Time-stamp TS (Txn-R). That is Txn-R gets the most recent version of the data before it, so reading-reading conflicts or reading -writing conflicts do not occur, and Txn-R are always succeeded.

b) Write-Only Transactions (Txn-W)

For this transaction type, the old data elements are not modified. Just a new version of data element will be created which is given by Txn-W as timestamp TS (Txn-W). So writing-writing conflicts do not occur and Txn-W will not be blocked by another transactions.

c) Update transactions (Txn-U)

This type of transaction not only reads the data elements but also writes a new version of data elements. So, writing-writing conflicts between update transactions most likely to be occurred Now, it is necessary to resolve the conflicts between the update transactions effectively to provide enhanced concurrency. The proposed method considers the contention of transactions in order to adopt best suited technique for that situation.

Contention is the number of transactions that are running in the system or waiting in concurrency control queue of Transaction manager to be executed. Conflicts among Update transactions (Txn-U) are resolved according to following rules:

If Contention is low, adopt OCC Sacrifice method. Allow all transactions (Txn-U) to be executed freely without any checking. When a transaction Txn-U_i reaches its validation stage, Txn-U_i checks for the conflicts with currently executing Txn-Us by means of the Read-set and Write-set of transactions. In the real-time database system Execution-Time (ET) of a transaction is predictable [1]. Let Conflict -Set (CS) be the set of Txn-Us that are conflicting with Txn-U_i. Now Txn-U_i restarts with rolling back its operations already performed, If ET(Txn- U_i) < $\sum_j$ ET(Txn- U_j) , For all Txn-U_j ϵ CS; otherwise Txn-U_i commits and For all Txn- U_j ϵ CS restart with updated data item.

If Contention is medium, adopt SCC method. When a transaction Txn-U_i begins to execute, it issues Exclusive Write (EW) lock on data object. If it completes the work with an data object D, it produces a new-image of that object D_n and converts EW lock into Speculative-Write (SPW) lock. After producing D_n, Txn-U_i allows any transaction Txn-U_j requesting for D to get speculative execution Txn-U_j begins speculative execution with new-image D_n and old image D_o. After the completion of its operation, Txn-U_j commits speculative execution with D_n or D_o according to the abort or commit of Txn-U_i respectively.

If Contention is high, adopt 2PL-HP that ensures more transactions to be completed within the deadline. The transaction priority P (Txn-U) is principally determined by deadline of the transactions. So as in Real time database system we can have the deadline of the transactions and take it as the priority for that transaction. That is for all Txn-U_i, Txn-U_j ϵ T, whenever deadline (Txn-U_i) > = deadline (Txn-U_j), then P (Txn-U_i) < = P(Txn-U_j), the higher-priority transaction will get the priority of execution. When a transaction Txn-U_j begins to execute, it tries to issue write lock on data object D. If D is already occupied by another transaction Txn-U_i , Txn-U_i have to sacrifice and restart. If P (Txn-U_j) > P(Txn-U_i) otherwise Txn-U_j waits for D to be free by the completion or abort of Txn-U_i.

III. Analysis of the Performance

Through the comparison testing between the new and traditional method, Fig.1 and Fig.2 shows how transactions different inter-arrival time affects the transactions restart. Analyzing the comparison figure Fig.1, It can be easily understood that, as the interval between transaction arrival increases, Contention decreases, the rate of restart of transaction becomes small due to the less opportunity of conflicts. But when the interval is not long, Contention increases the new method is considerably better than conventional one. The performance of real-time database systems has a fundamentally different target compared with the traditional database systems. The real-time database systems require the more number of transactions to be completed within the deadline of the transactions rather than the number of concurrent transaction for execution to maintain the largest concurrency. The new method makes read-only and write- only transactions never fail and avoids their unnecessary restart using multi-version method. It effectively saves the systems expense and improves the systems through put. As for the update transactions, the new technique makes the same data element to be operated by more transactions without interfering by another one by using suitable method to resolve conflicts, according to the contention of transactions in the system. This method can adaptively use the optimistic mechanism and speculation based mechanisms for the implementation of the resolution of conflicts and creates a new version of data elements to improve the concurrency degree of transaction and the amount of transactions completed before deadline. In summary, in different contention size, the new method can flexibly take advantage of the traditional concurrency control mechanism with multiple versions, OCC Sacrifice, Speculative Concurrency Control and 2PL-HP; it can get better the concurrency of the system, save effectively the expense of the system. Compared with the traditional concurrency control mechanisms, the improved one is better on performance.

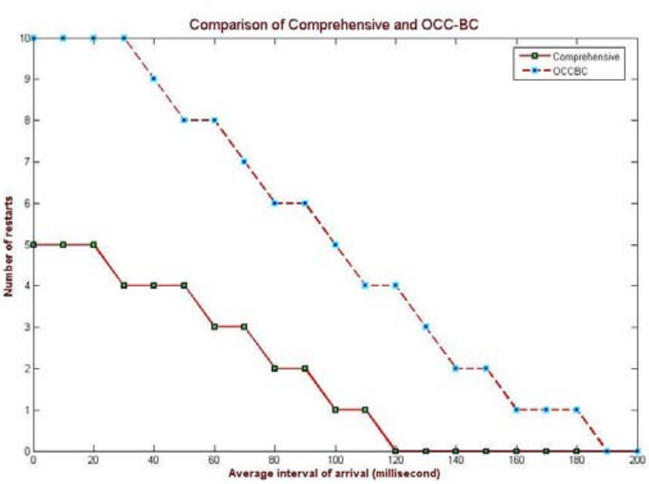


Figure 1 : Restarts vs. Average interval of arrival (For 10 Transactions)

IV. Conclusions

As the real-time database systems have a tight time constraints for the transactions as well as data, and the very well. So there is a demand of comprehensive traditional concurrency control mechanisms cannot satisfy their needs method to meet the requirements of the real time database system. By improving the concurrency control protocols, this paper has presented a comprehensive concurrency control technique that highly reduces the rate of abortion as well as considers the timing constraints of the transactions. With a strong self-adaptability, this method is able to use best suited method among different concurrency control mechanisms according to different situations of contention in the system. It can also effectively improve the performance of system providing higher concurrency considering deadline. The next step is to do further test and evaluation so that the other protocols can be justified with respect to comprehensive method that would be a good verification which is left as well as verification in the actual environment so as to refine and improve the algorithm.

References Références Referencias

1. S P. Wu, Z. Pang, "Research on the Improvement of the Concurrency Control Protocol for Real-Time Transaction", Proceedings of International Conference on Machine Vision and Human-Machine Interface, pp. 146-148, 2010.
2. K. M. Prakash Lingam, "Freezing as a correctness measure for multi-version time-stamp ordering protocol", Proceedings of2nd International Conference on Computer Engineering and Technology, vol.3, pp.312-316, 2010.
3. M. Hedayati, S. H. Kamali, R. Shakerian, M. Rahmani , "Evaluation of performance concurrency control algorithm for secure firm real-time database systems via simulation model", Proceeding of the International conference on Networking and Information Technology, pp. 260-264, 2010.
4. M. Laiho , F. Laux, "Implementing Optimistic Concurrency control for persistence Middleware using row version verification", Proceeding of the Second International Conference on Advances in databases, knowledge and Data Applications, pp. 45-50, 2010.
5. Rajib Mall, IISC, Bangalore "Real Time systems: Theory and Practice" , Ch-7, ISBN 10:8131771016
6. H. G. Molina, J. D. Ullman, J. Widom, "The Implementation of Database System", China Machine Press, 2002:369-377.
7. X. M. Yang, Y. X. Jun, "The comparative study on the implementation of concurrency control", Computer Application Research, 2006, (6):19-22.

8. C.Park, S.Park, S H. Son "Multi-version locking protocol with Freezing for secure Real-Time Database Systems", IEEE Transactions on Knowledge and Data Engineering, Vol. 14, no.5, pp.1141-1154, Oct -2002.
9. Fekete, D. Liarokapis, E. O'Neil, P. O'Neil and D. Shasha, "Making snapshot isolation serializable," ACM Trans. Database Syst., vol. 30,no. 2, pp. 492-528, 2005.
10. S. Jorwekar, A. Fekete, K. Ramamritham, and S. Sudarshan, "Automating the detection of snapshot isolation anomalies," Proceedings of the VLDB'07, 2007, pp. 1263-1274.
11. C. Park and S. Park, "Alternative Correctness Criteria for Multi-version Concurrency Control and a Locking Protocol via Freezing," Proc. Int'l Database Eng. and Applications Symp. pp. 73-81, Aug. 1997.
12. P.A. Bernstein and N. Goodman, "Multi-version Concurrency Control-Theory and Algorithms," ACM Trans. Database Systems, vol. 8, no. 4, pp. 465-483, Dec. 1983.
13. P. Bernstein, V. Hadzilacos, and N. Goodman, "Concurrency Control and Recovery in Database Systems," Addison-Wesley, 1987.
14. C. Park, S. Park, S H. Son, "Multi-version Locking Protocol with Freezing for Secure Real-Time Database Systems," IEEE Transactions on Knowledge and Data Engineering, vol. 14, no. 5, pp. 1141-1154, Oct. 2002.

Global Journal of Computer Science and Technology
Software & Data Engineering
Volume 13 Issue 2 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Survey of Techniques for Answering Top-k Queries

By Neethu C V & Rejimol Robinson R R
CT College of Engineering Trivandrum, India

Abstract - Top-k queries are useful in retrieving top-k records from a given set of records depending on the value of a function F on their attributes. Many techniques have been proposed in database literature for answering top-k queries. These are mainly categorized into three: Sorted-list based, layer based and View based. In first category, records are sorted along each dimension and then assigned a rank to each of the records using parallel scanning method. Threshold Algorithm (TA) and Fagin's Algorithm (FA) are the examples of sorted-list based category. Second category is layer based category, in which all the records are organized into layers such as in onion technique and robust indexing technique. Third category includes methods such as PREFER and LPTA (Linear Programming Adaptation of Threshold Algorithm) and processing is based on the materialized views.

Keywords : monotone functions, prefer, linearly optimally ordered set, convex hull.

GJCST-C Classification : H.2.3

A SURVEY OF TECHNIQUES FOR ANSWERING TOP-K QUERIES

Strictly as per the compliance and regulations of:

A Survey of Techniques for Answering Top-k Queries

Neethu C V [α] & Rejimol Robinson R R [σ]

Abstract - Top-k queries are useful in retrieving top-k records from a given set of records depending on the value of a function F on their attributes. Many techniques have been proposed in database literature for answering top-k queries. These are mainly categorized into three: Sorted-list based, layer based and View based. In first category, records are sorted along each dimension and then assigned a rank to each of the records using parallel scanning method. Threshold Algorithm (TA) and Fagin's Algorithm (FA) are the examples of sorted-list based category. Second category is layer based category, in which all the records are organized into layers such as in onion technique and robust indexing technique. Third category includes methods such as PREFER and LPTA (Linear Programming Adaptation of Threshold Algorithm) and processing is based on the materialized views.

Keywords : monotone functions, prefer, linearly optimally ordered set, convex hull.

I. Introduction

Top-k queries are intended for retrieving top-k records from the database which are subjected to minimization or maximization of the function F on the attributes of the relation. This kind of queries appears frequently in many applications such as college ranking, job ranking etc. Due to the popularity of top-k queries, many techniques have been proposed which are mainly includes sorted-list based, layer based and view based techniques.

a) Sorted-list based

Methods in this category sorts all records along each dimension and then assigned an overall grade to each of the records based on the sorted lists. For example, consider the example of college ranking. A student wants to join a college for doing graduation and he has some preferences based on the attributes like distance to the college, tuition fee and university under which college is working, performance of the college for previous four years etc. He then assigns grades to each of the attributes and sorted lists are created based on this assignment corresponding to each of the attributes. Then a list of colleges has retrieved based on their value for the query function. Here, the query function is a linear function in terms of the attributes of the records. FA and TA [1], [2], [3] are the two techniques included in this category.

Author α : Dept. of Computer Science & Engineering SCT College of Engineering Trivandrum, India.
E-mail : neethusureshbabu@gmail.com

b) Layer Based Category

The algorithms in this category organize all records into consecutive layers, such as Onion [4] and Robust Indexing Techniques [5]. The organization strategy is based on the common property among the records, such as the same convex hull layer in Onion [4]. Any top-k query can be answered by up to k layers of records. The Onion indexing is based on a geometric property of convex hull, which guarantees that the optimal value can always be found at one or more of its vertices.

The Onion indexing makes use of this property to construct convex hulls in layers with outer layers enclosing inner layers geometrically. A data record is indexed by its layer number or equivalently its depth in the layered convex hull. Queries with linear weightings issued at run time are evaluated from the outmost layer inwards. Onion indexing achieves orders of magnitude speedup against sequential linear scan when N is small compared to the cardinality of the set. The Onion technique also enables progressive retrieval, which processes and returns ranked results in a progressive manner. Furthermore, the proposed indexing can be extended into a hierarchical organization of data to accommodate both global and local queries.

Robust indexing [5] method is a kind layered technique for answering ranked queries. The layered indexing methods are less sensitive to the query weights. A key observation is that it may be beneficial to push a tuple as deeply as possible so that it has less chance to be touched in query execution. Motivated by this, a new criterion for sequentially layered indexing had been proposed: for any k, the number of tuples in top k layers is minimal in comparison with all the other layered alternatives. Since any top-k query can be answered by at most k layers, this proposal aims at minimizing the worst case performance on any top-k queries. Hence the proposed index is robust. While Onion and other layered techniques are sensitive to the query weights, this method, even though not optimal in some cases, has the best expected performance. Another appealing advantage of our proposal is that the top-k query processing can be seamlessly integrated into current commercial databases. Both Onion and other layered methods require the advanced query execution algorithms, which are not supported by many database query engines so far.

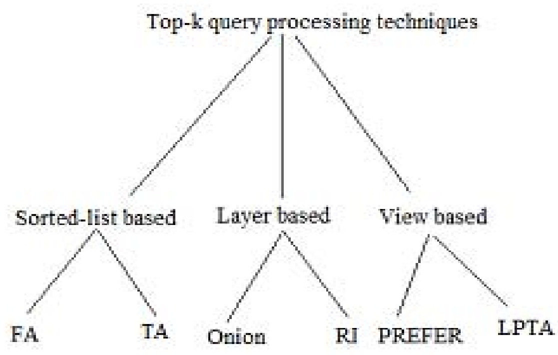


Figure 1 : Classification of Top-k query evaluation techniques

c) View based category

In view based techniques, the materialized views created from the relation can be used to answer top-k queries. PREFER [6] answers preference queries efficiently by using materialized views that have been preprocessed and stored. Queries with different weights will be first mapped to the pre-computed order and then answered by determining the lower bound value on that order. When the query weights are close to the pre-computed weights, the query can be answered extremely fast. Unfortunately, this method is very sensitive to weighting parameters. A reasonable derivation of the query weights (from the pre-computed weights) may severely deteriorate the query performance. PREFER is a layer on top of commercial relational databases and allows the efficient evaluation of multi parametric ranked queries. LPTA [7] is a linear programming adaptation of the classical TA algorithm to solve top-k query problem.

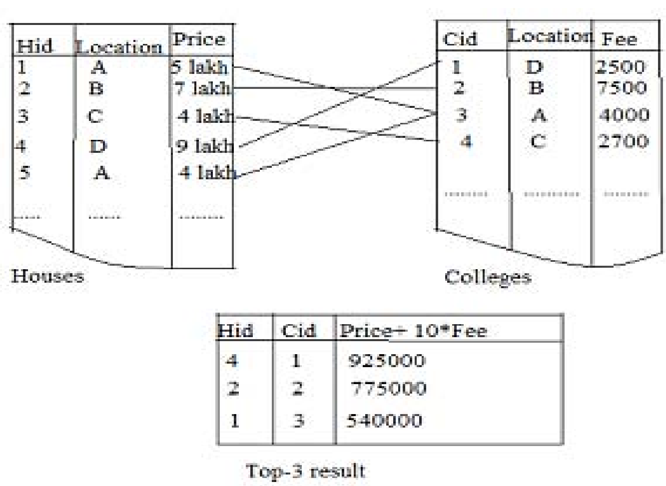


Figure 2 : Example of top-k query processing

II. Taxonomy of Processing Top-K Queries

Due to the high popularity of the top-k queries, various techniques have been proposed for solving such situations. Supporting efficient top-k query processing in database system is relatively recent and active line of research. In the following subsection, all the important techniques included in above explained categoris have been explored in detail.

a) Naïve Algorithm

To determine the top k objects, that is, k objects with the highest overall grades, the naive algorithm must access every object in the database, to find its grade under each attribute.

Step of the Naïve algorithm [1] are given below.

- If (x1, x2,…,xm) are the grades of object R under the m attributes, then compute T(x1,x2,…,xm) overall grade of object R.
- Sort the list of computed values.
- Return top k rows corresponding to the sorted list.

The main disadvantage of the Naïve algorithm is the large processing time when dealing with large databases.

b) Fagin's Algorithm

Fagin introduced an algorithm (''Fagin's Algorithm [1]'', or FA), which often does much better than the naive algorithm. In the case where the orderings in the sorted lists are probabilistically independent, FA finds the top k answers, over a database with N objects with arbitrarily high probability. This algorithm is implemented in Garlic, an experimental IBM middleware system.

- Do sorted access in parallel to each of the m sorted lists Li: Wait until there are at least k ''matches'', that is, wait until there is a set of at least k objects such that each of these objects has been seen in each of the m lists.
- For each object R that has been seen, do random access as needed to each of the lists Li to find the ith field xi of R:
- Compute the grade t(R)= t(x1,x2,….xm) for each object R that has been seen. Let Y be a set containing the k objects that have been seen with the highest grades (ties are broken arbitrarily). The output is then the graded set {(R, t(R)) | R€Y}.

Fagin shows that his algorithm is optimal with high probability in the worst case if the aggregation function is strict (so that, intuitively, we are dealing with a notion of conjunction), and if the orderings in the sorted lists are probabilistically independent. In fact, the access pattern of FA is oblivious to the choice of aggregation function, and so for each fixed database, the middleware cost of FA is exactly the same no matter what the aggregation function is. This is true even for a constant aggregation function; in this case, of course, there is a trivial algorithm that gives us the top k answers (any k objects will do) with O(1) middleware cost. So FA is not optimal in any sense for some monotone aggregation functions t: As a more interesting example, when the aggregation function is max (which is not strict), it is shown in that there is a simple algorithm that makes at most m*k sorted accesses and no random accesses that finds the top k answers. By contrast, the algorithm TA is instance optimal for every monotone aggregation function, under very weak assumptions.

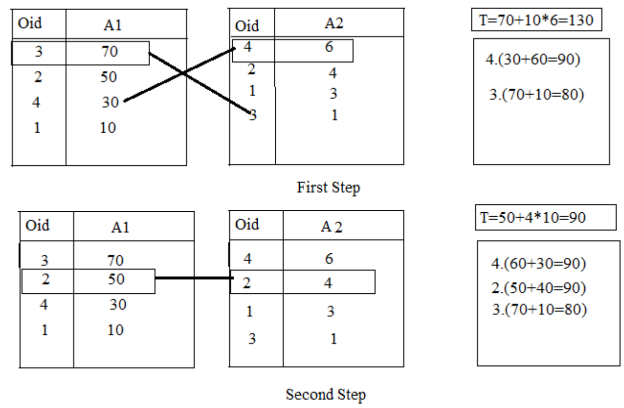
III. Threshold Algorithm

Even in the cases where FA is optimal, this optimality holds only in the worst case, with high probability. This leaves open the possibility that there are some algorithms that have much better middleware cost than FA over certain databases. The algorithm TA, which we now discuss, is such an algorithm.

- Do sorted access in parallel to each of the m sorted lists Li: As an object R is seen under sorted access in some list, do random accesses to the other lists to find the grade xi of object R in every list Li.
- Then compute the grade t(R) =t(x1,x2,…xm) of object R: If this grade is one of the k highest we have seen, then remember object R and its grade t(R).
- For each list Li, let xi be the grade of the last object seen under sorted access. Define the threshold value ψ to be t(x1, x2,….,xm). As soon as at least k objects have been seen whose grade is at least equal to ψ then halt.
- Let Y be a set containing the k objects that have been seen with the highest grades. The output is then the graded set {(R, t(R)) | R€Y}.

The algorithm scans multiple lists, representing different rankings of the same set of objects. An upper bound T is maintained for the overall score of unseen objects. The upper bound is computed by applying the scoring function to the partial scores of the last seen objects in different lists. Notice that the last seen objects in different lists could be different. The upper bound is updated every time a new object appears in one of the lists. The overall score of some seen object is computed by applying the scoring function to object's partial scores, obtained from different lists. To obtain such partial scores, each newly seen object in one of the lists is looked up in all other lists, and its scores are aggregated using the scoring function to obtain the overall score. All objects with total scores that are greater than or equal to T can be reported. The algorithm terminates after returning the kth output. Example 1 given below illustrates the processing of TA.

Example 1 (10): Consider two data sources containing same set of objects. Let A1 and A2 are the attributes in two data sources respectively. The Query function, F is defined as F=A1+10*A2. The working of TA is depicted in the following figure.



In the first step, retrieving the top object from each list, and probing the value of its other attribute value in the other list, results in revealing the exact scores for the top objects. The seen objects are buffered in the order of their scores. A threshold value, T, for the scores of unseen objects is computed by applying F to the last seen scores in both lists, which results in 70+6*10=130. Since both seen objects have scores less than T, no results can be reported. In the second step, T drops to 90, and objects 4 and 2 can be safely reported since its score is above T. The algorithm continues until k objects are reported, or sources are exhausted.

IV. Onion Technique

This technique comes under the layer based category and uses a special indexing structure for answering top-k queries. The Onion indexing is based on a geometric property of convex hull, which guarantees that the optimal value can always be found at one or more of its vertices. The Onion indexing makes use of this property to construct convex hulls in layers with outer layers enclosing inner layers geometrically. A data record is indexed by its layer number or equivalently its depth in the layered convex hull. Queries with linear weightings issued at run time are evaluated from the outmost layer inwards.

Basic idea of the onion technique is that partition the collection of d-dimensional data points into sets that are optimally linearly ordered. This property is used to construct convex hulls in layers with outer layers enclosing inner layers geometrically.

Definition 1: Optimally Linearly Ordered Set: A collection of sets{s1,s2,…,sn}are optimally linearly ordered sets if and only if a d-dimensional vector ā,

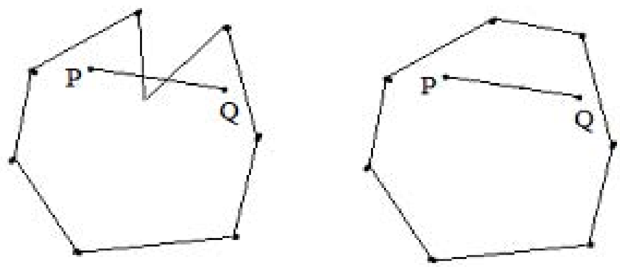
$$\exists\, \bar{o} \in s_i \text{ such that}$$

for every $\hat{c} \in s_{i+j}$, $j>0$, $\bar{a}^t \bar{o} > \bar{a}^t \hat{c}$ where $\bar{a}^t \bar{o}$ represents the inner product of two vectors.

Partitioning a set of data points into optimally linearly ordered sets is based on the following theorem.

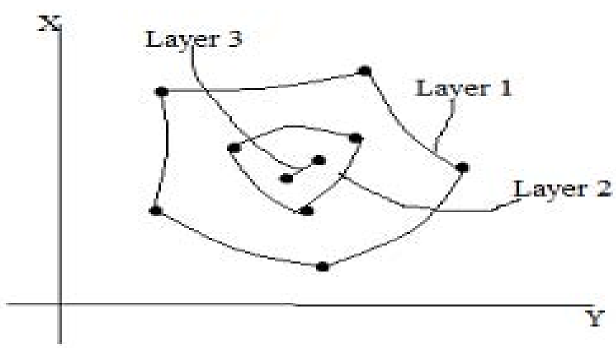
Theorem 1: Given a set of records R mapped to a d-dimensional space, and a linear maximization criterion, the maximum objective value is achieved at one or more vertices of the convex hull of R.

Definition 2 : A set S is convex if whenever two points P and Q are inside S, then the whole line segment PQ is also in S.



Non convex hull Convex hull

Procedure for index creation:
Step 1: Input a set of records R and iterate the following steps until size(R) become less than zero.
Step 2: Construct convex hull of the data records R.
Step 3: Store the records of hull vertices in set Vi.
Step 4: Assign records in set V to layer k.
Step 5: Set R=R-V and k=k+1.



This indexing structure can be used for query evaluation. Onion indexing achieves orders of magnitude speed up against sequential linear scan when N is small compared to the cardinality of the set. The Onion technique also enables progressive retrieval, which processes and returns ranked results in a progressive manner. Furthermore, the proposed indexing can be extended into a hierarchical organization of data to accommodate both global and local queries.

V. Robust Indexing Structure

This is an another layered indexing structure useful for the evaluation of top-k queries. The idea of multi-layer indexing has been also adopted by to provide robust indexing [5], [10] for top-k queries. Robustness is defined in terms of providing the best possible performance in worst case scenario, which is fully scanning the first k layers to find the top-k answers. The main idea is that if each object Oi is pushed to the deepest possible layer, its retrieval can be avoided if it is unnecessary. This is accomplished by searching for the minimum rank of each object oi in all linear scoring functions. Such rank represents the layer number, denoted l*(Oi), where object Oi is pushed to. For n objects having d scoring predicates, computing the exact layer numbers for all objects has a complexity of O(nd log n), which is an overkill when n or d are large. Approximation is used to reduce the computation cost. An approximate layer number, denoted l (Oi), is computed such that l (Oi) • l*(Oi), which ensures that no false positives are produced in the top-k query answer.

VI. Prefer

This is a view based evaluation of the top-k queries. Recent successful work in non-layered approaches includes the PREFER system [6],[10], where tuples are sorted by a pre-computed linear weighting configuration Users often need to optimize the selection of objects by appropriately weighting the importance of multiple object attributes. Such optimization problems appear often in operations research and applied mathematics as well as everyday life; e.g., a buyer may select a home as a weighted function of a number of attributes like its distance from office, its price, its area, etc.

The queries here use a weight function over a relation's attributes to derive a score for each tuple. Database systems cannot efficiently produce the top results of a preference query because they need to evaluate the weight function over all tuples of the relation. PREFER [6] answers preference queries efficiently by using materialized views that have been preprocessed and stored. Queries with different weights will be first mapped to the pre-computed order and then answered by determining the lower bound value on that order. When the query weights are close to the pre-computed weights, the query can be answered extremely fast. Unfortunately, this method is very sensitive to weighting parameters. A reasonable derivation of the query weights (from the pre-computed weights) may severely deteriorate the query performance. PREFER is a layer on top of commercial relational databases and allows the efficient evaluation of multi parametric ranked queries For example consider a database containing houses available for sale. The properties have attributes such as price, number of bedrooms, age, square feet, etc. For a user, the price of a property and the square feet area may be the most important issues, equally weighted in the final choice of a property, and the property's age may also be an important issue, but of lesser weight. The vast majority of e-commerce systems available for such applications do not help users in answering such queries, as they commonly order according to a single attribute. In these cases, preference queries have significant role and for PRFER system also.

VII. LPTA

Algorithm (LPTA)[7],[10] is another technique included in the view based category.It performs much better than
PREFER.

Problem 1: (Top-K Query Answer Using Views). Given a set U of views, and a query Q, obtain an answer to Q combining all the information conveyed by the views in U.

Consider a single relation R with m numeric attributes X1, X2,….Xm, and n tuples t1, . . . , tn. Let Domi = [lbi, ubi] be the domain of the ith attribute. Refer to table R as a base table. Each tuple t may be viewed as a numeric vector t = (t[1], t[2], . . . , t[m]). Each tuple is associated with a tuple-id (tid).Here consider top-k ranking queries, which can be expressed in SQL-like syntax: SELECT TOP [k] FROM R WHERE RangeQ ORDER BY Score Q. More abstractly, a ranking query may be expressed as a triple Q = (ScoreQ, k, Range Q), where Score Q(t) is a function that assigns a numeric score to any tuple t (the function does not necessarily involve all attributes of the table), and Range Q(t) is a Boolean function that defines a selection condition for the tuples of R in the form of a conjunction of range restrictions on Domi, i 2 {1, . . . ,m}. Each range restriction is of the form li ≤ Xi ≤ ui, I ϵ {1, . . . ,m} and the interval [li, ui] Domi. The semantics requires that the system retrieve the k tuples with the top scores satisfying the selection condition.

LPTA [7] is a linear programming adaptation of the classical TA algorithm to solve Problem 1.1 for the special case when views and queries are of the form V 0 = (Score V 0, n, *) and Q = (Score Q, k, *) respectively. Consider a relation with attributes X1, X2 and X3 as shown in Figure 1.3.2.1. Let views V1 and V2 have scoring functions f1, f2 respectively as shown in Figure 1.3.2.1 and consider a query Q = (f3, k, *).

The algorithm initializes the top-k buffer to empty. It then starts retrieving the tids from the views V1, V2 in a lock-step fashion, in the order of decreasing score (w.r.t. the view's scoring functions). For each tid read, the algorithm retrieves the corresponding tuple by random access on R, computes its score according to the query's scoring function f3, updates the top-k buffer to contain the top-k largest scores (according to the query's scoring function), and checks for the stopping condition as follows: After the dth iteration, let the last tuple read from view V1 be (tidd1, sd1) and from view V2 be (tidd2, sd2). Let the minimum score in the top-k buffer be topkmin. At this stage, the unseen tuples in the view have to satisfy the following inequalities (the domain of each attribute of R of Figure is [1, 100]).

						f1 = 2x1+5x2			f2 = x2+2x3	
R	tid	X1	X2	X3	V1	tid	score	V2	tid	score
						7	527		6	219
	1	82	1	59		6	299		4	202
	2	53	19	83		4	270		10	197
	3	29	1	2		8	246			
	4	80	22	90		2	201			
	5	28	8	87						
	6	12	55	82						
	7	16	99	42						
	8	18	42	67						
	9	42	1	23						
	10	23	21	88						

Figure 3 : Example of views

The following system of inequalities defines a Convex region in three dimensional space.

$$0 \leq X_1, X_2, X_3 \leq 100 \quad (1)$$
$$2X_1 + 5X_2 \leq s_d^1 \quad (2)$$
$$X2 + 2X3 \leq s_d^2 \quad (3)$$

This system of inequalities defines a convex region in three dimensional space. Let unseenmax be the solution to the linear program where we maximize the function f3 = 3X1 + 10X2 + 5X3 subject to these inequalities. It is easy to see that unseenmax represents the maximum possible score (with respect to the ranking query's scoring function) of any tuple not yet visited in the views. The algorithm terminates when the top-k buffer is full and unseenmax ≤ topkmin. Considering the example of given figure, the algorithm will proceed as follows;

First retrieve tid and conduct a random access to R to retrieve the full tuple and tid 6 from V2 accessing R again. The top-2 buffer contains the following pairs (tiddi, sdi) {(7, 1248), (6, 996)}. The solution to the linear program with s1q= 527 and s2d = 219 yields an unseenmax =1338 > topkmax = 1248 and the algorithm conducts one more iteration.This time we access tid 6 from V1 and tid 4 from V2. The top-2 buffer remains unchanged and the linear program is solved one more time using sd1 = 299 and sd 2 = 202. This time, unseen max= 953.5 < topkmax = 1248 and the algorithm terminates. Thus, in total LPTA conducts two sequential and two random accesses per view. In contrast, the TA algorithm executed on R of Figure 1 will identify the correct top-2 results after 12 sorted and 12 random accesses in total. The performance advantage of LPTA is evident.

Algorithm LPTA(U,Q)

U={V1,…,V_r}//set of views
Q=($Score_Q$,k,*)//Query
Topk-Buffer={}
$topk_{min}$ =0
for d=1 to n do
for all views Vi(1≤i≤r) in block-step do
Let(tid^i_d, s^i_d) be the d-th item in Vi
//Update top-k buffer
Let t^i_d =RandomAccess(tid^i_d)
if $Score_Q$(t^i_d)>$topk_{min}$ **then**
if(|topk-Buffer|=k) **then**
Remove min score tuple from topk-Buffer
end if
Add(tid^i_d,$Score_Q$(tid^i_d)) to topk-buffer
$Topk_{min}$=min score of topk-Buffer
end if
// Checking stopping conditions by solving LP
Let $Unseen_{max}$ =convex region defined by
lb_j≤X_j≤ub_j for every 1≤j≤m
$Score_{vj}$ ≤s^j_d for every 1≤j≤r
Compute $Unseen_{max}$ =$max_{tcunseen}${$Score_Q$(t)}
If(|topk-Buffer|=k) and ($Unseen_{max}$≤$topk_{min}$)then
Return topk-Buffer
end if
end for
end for

VIII. Comparison of Different Techniques

This section includes comparison of different techniques employed in the top-k query evaluation. The comparison is performed based on the three important criteria which are ranking function, ranking model and data access operation involved in the different techniques. The ranking function can be generic or monotone. Most of the current top processing techniques assume monotone ranking functions since they fit in many practical scenarios, and have appealing properties allowing for efficient top-k processing. But Few recent techniques address top-k queries in the context of constrained function optimization. The ranking function in this case is allowed to take a generic form.

	Ranking Function		Ranking Model		Data Access	
	Generic	Monotone	Top-k join	Top-k selection	Sorted	Random
PREFER		•	•	•	N/A	N/A
Onion Technique		•		•	N/A	N/A
TA		•		•	•	•
FA		•		•	•	•
Naive	•			•	•	

Another criteria is ranking model. It can be top-k join or top-k selection. In top-k selection model, the scores are assumed to be attached to base tuples. A top-k selection query is required to report the k tuples with the highest scores. Scores might not be readily available since they could be the outcome of some user-de Consider a set of relations R1 ,….,Rn. A top-k join query joins R1,…,Rn, and returns the k join results with the largest combined scores. The combined score of each join result is computed according to some function F(p1,…., pm), where p1,….,pm are scoring predicates defined over the join results. Fined scoring function that aggregates information coming from different tuple attributes. Third criteria is data access which can be sorted access or random access. In sorted access, Object R has the Ith highest grade in the ith list, then I sorted accesses to the ith list are required to see the grade under sorted access and in random access, grade of object R in the ith list obtains it in one random access.

IX. Conclusion

A surevey of top-k query processing techniques based on the different criterias have done. For this purpose, a detailed analysis of different techniques included in three important categories like sorted-list based category, layer based category and view based category have explored.

References Références Referencias

1. R. Fagin, A. Lotem, and M. Naor, “Optimal Aggregation Algorithms for Middleware,” Proc. Symp. Principles of Database Systems (PODS), 2001.
2. S. Nepal and M.V. Ramakrishna, “Query Processing Issues in Image (Multimedia) Databases,” Proc. 15th Int’l Conf. Data Eng. (ICDE), 1999.
3. U. Guntzer, W.T. Balke, and W. Kiebling, “Optimizing Multi-Feature Queries for Image Databases,” Proc. Int’l Conf. Very Large Data Bases (VLDB), 2000.
4. Y.-C. Chang, L.D. Bergman, V. Castelli, C.S. Li, M.L. Lo, and J.R. Smith, “The Onion Technique: Indexing for Linear Optimization Queries,” Proc. ACM SIGMOD, 2000.
5. D. Xin, C. Chen, and J. Han, “Towards Robust Indexing for Ranked Queries,” Proc. Int’l Conf. Very Large Data Bases (VLDB), 2006.
6. Hristidis, N. Koudas, and Y. Papakonstantinou, “Prefer: A System for the Efficient Execution of Multi -Parametric Ranked Queries,” Proc. ACM SIGMOD, 2001.
7. G. Das, D. Gunopulos, N. Koudas, and D. Tsirogiannis, “Answering Top-K Queries Using Views,” Proc. Int’l Conf. Very Large Data Bases (VLDB), 2006.
8. S. Bo¨ rzso¨ nyi, D. Kossmann, and K. Stocker,

"The Skyline Operator," Proc. 17th Int'l Conf. Data Eng. (ICDE), 2001.
9. D. Papadias, Y. Tao, G. Fu, and B. Seeger, "An Optimal and Progressive Algorithm for Skyline Queries," Proc. ACM SIGMOD, 2003.
10. Ihab F. Ilyas, George Beskales And Mohamed A. Soliman, "A survey of top-k query processing technique in relational database systems" University of Waterloo, Support was provided in part by the Natural Sciences and Engineering Research Council of Canada 2011.
11. D. Kossmann, F. Ramsak, and S. Rost, "Shooting Stars in the Sky: An Online Algorithm for Skyline Queries," Proc. Int'l Conf. Very Large Data Bases (VLDB), 2002.
12. T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein, Introduction to Algorithms. MIT Press, 2001.
13. D. Campbell and R. Nagahisa, "A Foundation for Pareto Aggregation," J. Economical Theory, vol. 64, pp. 277-285, 1994.
14. M. Voorneveld, "Characterization of Pareto Dominance," Operations Research Letters, vol. 32, no. 3, pp. 7-11, 2003.
15. C. Li, B.C. Ooi, A.K.H. Tung, and S. Wang, "DADA: A Data Cube for Dominant Relationship Analysis," Proc. ACM SIGMOD, 2006.
16. Y. Tao, V. Hristidis, D. Papadias, and Y. Papakonstantinou, "Branch-and-Bound Processing of Ranked Queries," Information Systems, vol. 32, no. 3, pp. 424-445, 2007.
17. R.J. Lipton, J.F. Naughton, and D.A. Schneider, "Practical Selectivity Estimation through Adaptive Sampling," Proc. ACM SIGMOD, 1990.
18. R.J. Lipton and J.F. Naughton, "Query Size Estimation by Adaptive Sampling," Proc. Symp. Principles of Database Systems (PODS), 1990.
19. S. Chaudhuri, N.N. Dalvi, and R. Kaushik, "Robust Cardinality and Cost Estimation for Skyline Operator," Proc. 22nd Int'l Conf. Data Eng. (ICDE), 2006.
20. C.B. Barber, D.P. Dobkin, and H. Huhdanpaa, "The QuickhullAlgorithm for Convex Hulls," ACM Trans. Math. Software, vol. 22, pp. 469-483, 1996.
21. D. Xin, J. Han, H. Cheng, and X. Li, "Answering Top-k Queries with Multi-Dimensional Selections: The Ranking Cube Approach," Proc. Int'l Conf. Very Large Data Bases (VLDB), 2006.
22. S. Nepal and M. Ramakrishna, "Query Processing Issues in Image (Multimedia) Databases," Proc. 15th Int'l Conf. Data Eng. (ICDE), 1999.
23. M. Li and Y. Liu, "Iso-Map: Energy-Efficient Contour Mapping in Wireless Sensor Networks," IEEE Trans. Knowledge and Data Eng., vol. 22, no. 5, pp. 699-710, May 2010.
24. H. Bast, D. Majumdar, R. Schenkel, M. Theobald, and G. Weikum, "IO-Top-k: Index-Access Optimized Top-k Query Processing," Proc. Int'l Conf. Very Large Data Bases (VLDB), 2006.
25. N. Mamoulis, K.H. Cheng, M.L. Yiu, and D.W. Cheung, "Efficient Aggregation of Ranked Inputs," Proc. 22nd Int'l Conf. Data Eng. (ICDE), 2006.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Software & Data Engineering
Volume 13 Issue 2 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Improved Algorithm for Frequent Item sets Mining Based on Apriori and FP-Tree

By Sujatha Dandu, B.L.Deekshatulu & Priti Chandra

Aurora's Technological and Research Institute, Hyderabad, India

Abstract - Frequent itemset mining plays an important role in association rule mining. The Apriori & FP-growth algorithms are the most famous algorithms which have their own shortcomings such as space complexity of the former and time complexity of the latter. Many existing algorithms are almost improved based on the two algorithms and one such is APFT [11], which combines the Apriori algorithm [1] and FP-tree structure of FP-growth algorithm [7]. The advantage of APFT is that it doesn't generate conditional & sub conditional patterns of the tree recursively and the results of the experiment show that it works fasts than Apriori and almost as fast as FP-growth. We have proposed to go one step further & modify the APFT to include correlated items & trim the non correlated itemsets. This additional feature optimizes the FP-tree & removes loosely associated items from the frequent itemsets. We choose to call this method as APFTC method which is APFT with correlation.

Keywords : data mining, correlation, correlation coefficient.

GJCST-C Classification : H.3.3

IMPROVED ALGORITHM FOR FREQUENT ITEM SETS MINING BASED ON APRIORI AND FP-TREE

Strictly as per the compliance and regulations of:

Improved Algorithm for Frequent Item sets Mining Based on Apriori and FP-Tree

Sujatha Dandu [α], B.L.Deekshatulu [σ] & Priti Chandra [ρ]

Abstract - Frequent itemset mining plays an important role in association rule mining. The Apriori & FP-growth algorithms are the most famous algorithms which have their own shortcomings such as space complexity of the former and time complexity of the latter. Many existing algorithms are almost improved based on the two algorithms and one such is APFT [11], which combines the Apriori algorithm [1] and FP-tree structure of FP-growth algorithm [7]. The advantage of APFT is that it doesn't generate conditional & sub conditional patterns of the tree recursively and the results of the experiment show that it works fasts than Apriori and almost as fast as FP-growth. We have proposed to go one step further & modify the APFT to include correlated items & trim the non correlated itemsets. This additional feature optimizes the FP-tree & removes loosely associated items from the frequent itemsets. We choose to call this method as APFTC method which is APFT with correlation.

Keywords : data mining, correlation, correlation coefficient.

I. Introduction

Association rules are if/then statements that help uncover relationships between seemingly unrelated data in a relational database. An association rule has two parts, an antecedent (if) and a consequent (then). An antecedent is an item found in the data. A consequent is an item that is found in combination with the antecedent.

Association rules are created by analyzing data for frequent if/then patterns and using the criteria of support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the database. Confidence indicates the number of times the if/then statements have been found to be true. Mining association rules purely on the basis of minimum support may not always give interesting relationships between the item sets.

Consider a case where in a sample set of 100 transactions, item A with support SA = 50 & item B with support SB = 50 have a combined support of SAB = 5. If the minimum support threshold is 5, it would appear as if A and B are frequent item sets because they satisfy the minimum support criteria. The drawback of this method is that only 10% of all A and 10% of all B are involved in association rules together. So the relationship between A and B cannot be of much use even though they occur together more than the support value. A new method is required wherein we measure not only the support but also the confidence of B occurring when A occurs & vice versa. This way we can make sure that the interestingness of the rules is preserved. The concept of correlation is introduced in order to filter the result from the association rules that not only satisfy the minimum support criteria but also have a linear relationship amongst them. This approach combines the concepts of FP-growths tree generation technique with Apriori's candidate generation step along with a correlation condition so as to improve the interestingness of rules as well as to optimize the space and time consumption.

II. Existing System

The concept of frequent itemset was first introduced by Agarwal et al in 1993.Two basic frequent itemset mining methodologies: Apriori & FP-growth, and their extensions, are introduced. Agarwal and Srikanth [2] observed an interesting downward closure property which states that: A k-itemset is frequent only if all of its sub-itemsets are frequent. It generates candidate itemset of length k from itemset of length k-1. Since the Apriori algorithm was proposed, there have been extensive studies on the improvements of Apriori, eg. partitioning technique[3], sampling approach[4], dynamic itemset counting [6], incremental mining [5] & so on.

Apriori, while historically significant, suffers from(1) generating a huge number of candidate sets, and (2) repeatedly scanning the database. Han et al [7] derived an FP-growth method, based on FP-tree. The first scan of the database derives a list of frequent items in which items in the frequency descending order are compressed into a frequent-pattern tree or FP-tree. The FP-tree is mined to generate itemsets. There are many alternatives and extensions to the FP-growth approach, including depth first generation of frequent itemset [8]; H-mine, by [9] which explores a hyper structure mining of frequent patterns; and an array-based implementation of prefix tree structure for efficient pattern growth mining [10]. To overcome the limitation of the two approaches a new method named APFT [11] was proposed. The APFT algorithm has two steps: first it constructs an FP-tree & then second mines the frequent items using Apriori

Author α : Aurora's Technological and Research Institute, Hyderabad, India. E-mail : sujatha.dandu@gmail.com
Author σ : Distinguished Fellow, IDRBT, Hyderabad, India. E-mail : deekshatulu@hotmail.com
Author ρ : Scientist, Advanced System Laboratory, Hyderabad, India. E-mail : priti_murali@yahoo.com

algorithm. [The results of the experiment show that it works faster than Apriori and almost as fast as FP-growth]. Extending this approach, we have introduced APFTC, which includes the concept of correlation to filter (reduce) the association rules that not only satisfy the minimum support but also have liner relationships among them. The computational results verify the good performance of APFTC algorithm.

III. Proposed System

a) Correlation Concept

The concept of correlation can be extended to transaction databases with the following modifications. An item 'a' is said to be correlated with item 'b' if it satisfies the following conditions:

P (ab) > P(a)P(b)
Here
P(ab) = probability of items 'a' and 'b' occurring together in the transaction database i.e. the number of transactions in which both 'a' and 'b' occur together/total number of transactions.
P(a) =The number of transactions in which 'a' occurs/total transactions.
P(b) = The number of transactions in which 'b' occurs/total transactions
Therefore the formula essentially represents

Observed probability > Expected Probability

This condition is said to be positive correlation between items 'a' and 'b'.

b) APFTC

The idea of correlation is introduced at step 5 of the APFT algorithm. We are basically deriving the frequent itemsets of size 2 at this step, so it is only appropriate to introduce the idea of correlation here. There is another change to the algorithm where we can calculate the support of each branch at the time of construction of calculation N Table itself instead of traversing the tree again later. This is a more economical way of calculating support than the one suggested in the original paper where repeated traversal of the tree is necessary for support calculation.

Algorithm APFT [ ]
Input: FP-tree, minimum support threshold ϶
Output: all frequent itemset L
L = L1;
for each item Ii in header table, in top down order
LIi = Apriori-mining (Ii);
Return L = {LULi1ULi2U…Lin};

Pseudo code Apriori-mining (I i)

1. Find item p in the header table which has the same name with Ii (
2. q = p.tablelink;
3. while q is not null
4. for each node qi != root on the prefix path of q
5. if N Table has a entry N such that N. Item-name = q i. item-name
6. N. Item-support = N. Item-support + q.count;
7. else
8. add an entry N to the N Table;
9. N.Item-name = q i. item-name;
10. N.Item-support = q.count;
11. q = q.tablelink;
12. k = 1;
13. Fk = {j | j"NTable)j.Item-support*minsup}
14. repeat
15. k = k + 1;
16. Ck = apriori-gen(Fk-1) ;
17. q = p.tablelink;
18. while q is not null
19. find prefix path t of q
20. Ct = subset(Ck, t);
21. for each c"+,t
22. c.support = c.support + q.count;
23. q = q.tablelink;
24. Fk = {c | c"+k,),c.support *,-./012}
25. until Fk = &
26. return LI i = Ii U F1 U F2 U…U Fk // Generate

Introducing correlation coefficient at line4, we continue for each node qi != root on the prefix path of q All Paths of Tree[i].add (qi.item-name);//AllPathsOfTree is an array of all paths from q to root. if NTable has a entry N such that N.Item-name= q i.item-name.

N.Item-support = N.Item-support + q.count; else

Check For Correlation Between qi and q
PA=Map Support (q.itemName);
PB=Map Support (qi.item-name);
P(AB)= q.count;
If(P(AB)>P(A)P(B)/transaction Count)
. add an entry N to the NTable;
. N.Item-name = q i. item-name;
. N.Item-support = q.count;

All Paths of Tree [i].support=q.count;//here we have the individual path //and its support stored to be used //later
q = q.tablelink;

c) Example

We follow an example of a simple database with 6 transactions as shown below.

Transaction Database
1 3 4
2 3 5
1 2 3 5
2 5
1 2 3 5

i. *FPTREE*

The fp-tree construction is a fairly straight forward procedure which follows from the above transaction database to the tree structure shown below

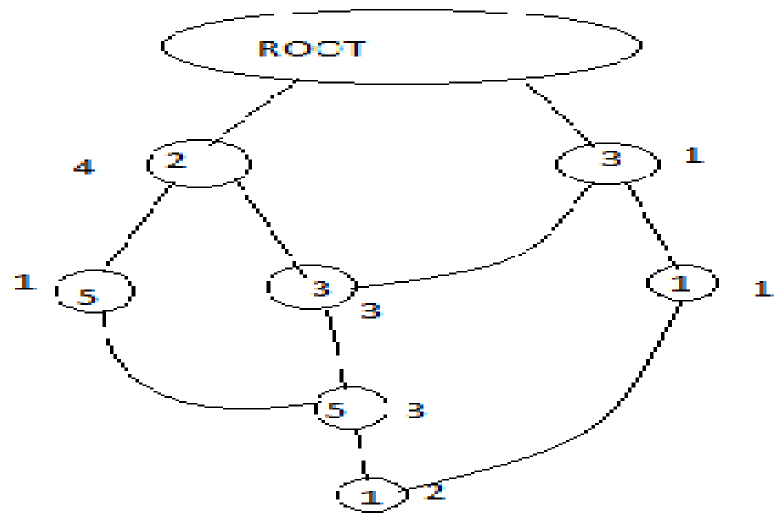


Figure 1 : FP-TREE

Once the tree has been constructed we proceed with the APFT algorithm with construction of an N Table for each of the nodes. We start of with the node 4 which is at the bottom of the header table for the given fp tree. Let us take the minimum support value as 2 for this example.

ii. *Ntable For Node 5*

Calculations for Correlation with item 5:

P (5, 2) =4/5=0.8
P (5) P (2)=4/5*4/5=0.64
Hence P (5, 2)>P (5) P (2)

P (5, 3) =3/5=0.6
P (5) P (3)=4/5*4/5=0.64
Hence P (5, 3) <P (5) P (3)

NODE	SUPPORT
2	4

OUTPUT
5 2:4

As shown in the figure we use the Ntable along with Apriori's candidate generation step to successively generate supersets of the the smaller itemsets and then perform the pruning step by calculating support by scanning the tree paths instead of scanning the entire database as is the case with apriori.

The candidate support calculation procedure is as shown in the diagram below each path from the node to the root is stored with that path's support.

IV. Results

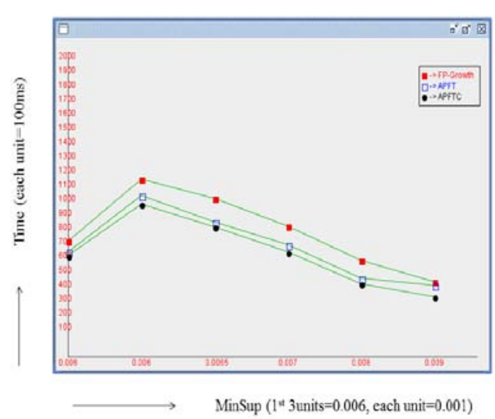
a) Datasets

The following datasets have been used in the experiment as inputs

DATASETS	Items	Max|t|	Avg|t|	Size
T10I4D100K	870	29	10	100000
Mushroom	119	23	23	8124

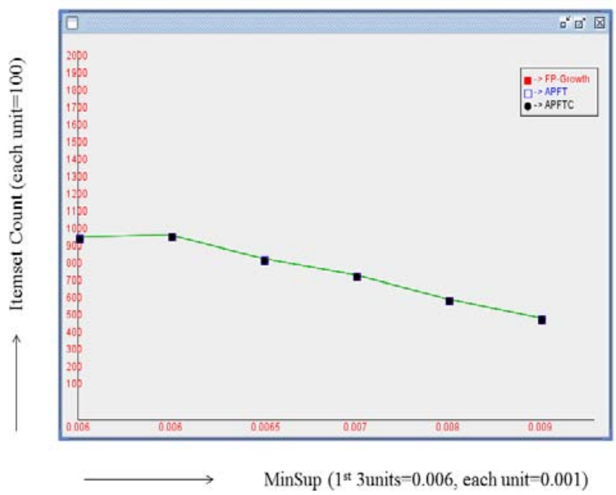
The algorithm APFTC is reportedly working efficiently and in many cases, it's much faster than FP-Growth. The results are found to be more interesting than association rules mined by FP-Growth although they are the subsets of itemsets mined by FP-Growth.

Figure 1 : Minimum Support vs. Time



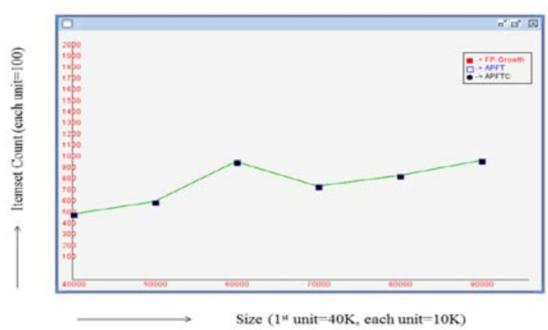
The above graph shows the nature of the three algorithms with varied minimum support. It can be observed that APFTC has completed the task in very less time when compared to the other two algorithms.

Figure 2 : Minimum Support vs. Count



The above graph shows the number of itemsets generated with respect to varying minimum support. Supporting our idea, APFTC has generated equal to the number of itemsets which are highly correlated when compared to the other algorithms.

Figure 3 : Size vs. Count



The above graphs gives the itemsets generated with respect to varying size. The itemsets generated by APFTC are equal to or less than those generated by the other two algorithms in some cases.

It can be concluded from the above results that APFTC performs as expected proving to be efficient in time consumed and also in retrieving the most correlated itemsets.

References Références Referencias

1. Data Mining – Concepts and Techniques – Jia Wei Han and Micheline Kamber
2. SPMF – A Sequential Pattern Mining Framework – Philippe Fournier-Viger
3. Contrasting Correlations by an Efficient Double-Clique Condition - Aixiang Li, Makoto Haraguchi, and Yoshiaki Okubo
4. R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. InProc.1993 ACM-SIGMOD Int. Conf. Management of Data, Washington, D.C., May 1993, pp 207–216.
5. R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In VLDBY94, pp. 487-499.
6. J.S. Park, M.S. Chen, and P.S. Yu. An effective hash- based algorithm for mining association rules. InSIGMOD1995, pp 175-186.
7. J. S. Park, M. S. Chen, and P. S. Yu. An effective hash- based algorithm for mining association rules. Proceedings of ACM SIGMOD International Conference on Management of Data, San Jose, CA, 1995, pp175-186.
8. Savasere, E. Omiecinski, and S. Navathe. An efficient algorithm for mining association rules in large databases. Proceedings of the 21st International Conference on Very large Database, 1995.
9. J. Han, J. Pei, and Y. Yin. Mining Frequent Patterns without Candidate Generation (PDF), (Slides), Proc. 2000ACM-SIGMOD Int. May 2000.
10. AgarwalR, AggarwalC, PrasadVVV.A tree projection algorithm for generation of frequent itemsets. In Journal of Parallel and Distributed Computing (Special Issue on High Performance Data Mining), 2000.
11. J. Pei, J. Han, and H. Lu. Hmine: Hyper-structure mining of frequent patterns in large databases. In ICDM, 2001, pp441–448.

Global Journal of Computer Science and Technology
Software & Data Engineering
Volume 13 Issue 2 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Efficient Concurrency Control Technique for Mobile Database Environment

By Md. Anisur Rahman

Dhaka University of Engineering & Technology Gazipur, Bangladesh

Abstract - Day by day, wireless networking technology and mobile computing devices are becoming more popular for their mobility as well as great functionality. Now it is an extremely growing demand to process mobile transactions in mobile databases that allow mobile users to access and operate data anytime and anywhere, irrespective of their physical positions. Information is shared among multiple clients and can be modified by each client independently. However, for the assurance of timely access and correct results in concurrent mobile transactions, concurrency control techniques (CCT) happen to be very difficult. Due to the properties of Mobile databases e.g. inadequate bandwidth, small processing capability, unreliable communication, mobility etc. existing mobile database CCTs cannot employ effectively. With the client-server model, applying common classic pessimistic techniques of concurrency control (like 2PL) in mobile database leads to long duration Blocking and increasing waiting time of transactions. Because of high rate of aborting transactions, optimistic techniques aren`t appropriate in mobile database as well. This paper discusses the issues that need to be addressed when designing a CCT technique for Mobile databases, analyses the existing scheme of CCT and justify their performance limitations. A modified optimistic concurrency control scheme is proposed which is based on the number of data items cached, amount of execution time and current load of the database server. Experimental results show performance benefits, such as increase in average response time and decrease in waiting time of the transactions.

Keywords : mobile database, optimistic concurrency control, mobile host, fixed host, base station, client-server.

GJCST-C Classification : H.2.4

AN EFFICIENT CONCURRENCY CONTROL TECHNIQUE FOR MOBILE DATABASE ENVIRONMENT

Strictly as per the compliance and regulations of:

An Efficient Concurrency Control Technique for Mobile Database Environment

Md. Anisur Rahman

Abstract - Day by day, wireless networking technology and mobile computing devices are becoming more popular for their mobility as well as great functionality. Now it is an extremely growing demand to process mobile transactions in mobile databases that allow mobile users to access and operate data anytime and anywhere, irrespective of their physical positions. Information is shared among multiple clients and can be modified by each client independently. However, for the assurance of timely access and correct results in concurrent mobile transactions, concurrency control techniques (CCT) happen to be very difficult. Due to the properties of Mobile databases e.g. inadequate bandwidth, small processing capability, unreliable communication, mobility etc. existing mobile database CCTs cannot employ effectively. With the client-server model, applying common classic pessimistic techniques of concurrency control (like 2PL) in mobile database leads to long duration Blocking and increasing waiting time of transactions. Because of high rate of aborting transactions, optimistic techniques aren`t appropriate in mobile database as well. This paper discusses the issues that need to be addressed when designing a CCT technique for Mobile databases, analyses the existing scheme of CCT and justify their performance limitations. A modified optimistic concurrency control scheme is proposed which is based on the number of data items cached, amount of execution time and current load of the database server. Experimental results show performance benefits, such as increase in average response time and decrease in waiting time of the transactions.

Keywords : mobile database, optimistic concurrency control, mobile host, fixed host, base station, client-server.

I. Introduction

In the modern decade, Because of having enhanced processing capability, easy and swift access to information as well as reduction in prices, mobile devices have achieved vast use in data processing services like mobile banking, traffic control, e-commerce, money transfers etc.

In these application and communication process of mobile devices client-server, peer to peer, and ad-hoc architectures are most recognized architectures of mobile database system. Fig. 1 displays client-server architecture. Mobile clients/Host (MH) e.g. Palmtops, Laptops, PDA's, Cellular phones etc. and fix host (FH) e.g. Terminals, desktops, servers etc and mobile base stations (BS) are three most important elements of this model. In client-server model, mobile clients are connected to fixed host through mobile base station. At present, using wireless 802.11 protocols, it is possible to have a synchronized and fast connection between mobile client and server. But, properties of mobile devices still impose some limitations upon assessment and data processing in mobile environment. Therefore, for constant and continuous processing, required data are copied partially from server to mobile client.

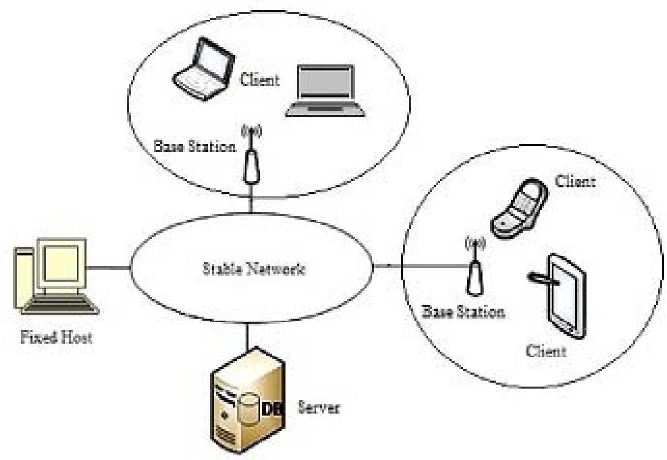


Figure 1 : Mobile Database Environment

In a mobile database environment, each node has an area of influence called cell, only within which others can receive its transmissions. In a distributed client-server mobile database system, not only clients but also servers are mobile, wireless and battery-powered. Mobile database systems are of interest because, it can be easily deployed in a short time, and end users can access and operate data anytime and anywhere. Examples of mobile database applications include law enforcement operations, automated battlefield applications, natural disaster recovery situations where the communication infrastructures have been destroyed, self-organizing sensor networks for data collecting, and interactive lectures or conferences for data exchange without pre-installed infrastructures.

However, the flexibility and convenience in Mobile Databases raise new issues when performing concurrency control (CC). CC is the activity of preventing transactions from destroying the consistency

Author : Department of CSE, Dhaka University of Engineering & Technology Gazipur, Bangladesh.
E-mail : ansiur.rahman.duet@gmail.com

of databases while allowing them to run concurrently, so that the throughput and resource utilization of database systems are improved and waiting time of concurrent transactions is reduced. When designing a CC algorithm the factors that should be taken into considerations are mobility, low bandwidth, multi-hop communication, limited battery power, limited storage, frequent disconnections, wireless communication delay, less processing power, frequent disconnections and unbounded disconnection time etc. When the execution is prolonged, the probability of conflicts with other transactions becomes higher and, consequently, transactions are likely blocked if a pessimistic CC method applies or restarted if an optimistic CC method is in use. The existing CC techniques for traditional mobile network databases, in which only clients are mobile and battery-powered, cannot be directly, because of a number of factors related to its architecture, availability and sharing of hardware and software resources, distribution of data, and mobile client's processing capability.

In this paper we proposed a new scheme that is reduces number of abortion as well as minimize the connection time of client and server by enhancing concurrency.

II. Related Works

Due to the boundaries, restrictions and specific properties of mobile devices cause traditional concurrency control techniques e.g. 2PL to exhibit a reduced amount of effectiveness in mobile database. A variety of researches have done in this field. These researches introduce new techniques of concurrency control or adapt the existing techniques of concurrency control with the requirements of mobile environment.

Because of the inherent limitations of mobile devices, constant and uninterrupted connection between client and server is not achievable over the transaction period. In locking-based pessimistic protocols, continuity of connection of mobile client with server is compulsory. If at the period of transaction, the connection is interrupted, the possibility of infinite blocking of transactions and deadlock is appeared. So, based on time out, many researchers have been done regarding methods of non-exclusive locks. In these techniques, if transaction doesn't end within the expected time, locks get caught by transaction will be free.

References [9, 10] showed a technique of locking with non-exclusive lock based on time out. In reference [10], a dynamic timer was used to solve the problem of transactions` blocking. In this method, transactions should be finished in particular time, otherwise they would abort. In reference [11], for increasing the efficiency of [10],transactions which don't finish in specified time and are near to final of transaction, wouldn't be aborted. They are allowed to continue execution for specified time. In methods basing on timer (time out) problems like long connection of mobile client with Server, estimated the amount of timer, and the length of transaction, do exist. Because of wrong estimation of the required time for completion of transactions, these problems could result in incorrect abortion of transactions. This suffers from the problem of frequent rollbacks due to regular expiry of the timer and wastage of computation. A protocol based on AVI is proposed in [12]. This method has still the problem of transaction blocking and computational overhead. This problem has been mentioned in reference [13]. Multi-version concurrency control based on MV2PL protocol and timestamp being introduced in [14] are in fact an extension of method [15].

In [17, 18], combination of optimistic and pessimistic is performed according to the semantic of operators.

In [17], changing two phases locking (2PL) and considering the semantic of transaction`s operators, are executed for increasing the concurrency degree of transactions. If conflicted operators are compatible semantically, then, they can choose a resource simultaneously. In this method, if compatible transactions lock a resource, there would be a resource scarcity for incompatible transaction. Moreover, there is the possibility of high rate of abortion through reconciliation process.

In reference [13] optimistic method is utilized. At the time of completion of transactions, updated data are sent to other mobile transactions using these data. This transfer is in multi-cast status. This technique reduces abortion rate of transactions.

Reference [19] puts emphasis on asymmetry of connecting bandwidth between mobile client and server. In [19] optimistic method is under consideration in which in the time of data updating, timestamp of transactions is set dynamically and data is broadcast for mobile clients.

Optimistic concurrency control based on timestamp ordering in [20] is adapted for broadcast environments. Also the introduced method in [21] is suitable for broadcast environments.

[3] OPCOT algorithm is introduced by assigned timestamp to operators and based on optimistic method according to concept of commitment ordering schedulers. So, reducing the waiting time as well as abortion rate, providing high concurrency, secure from deadlock and starvation, reducing computational are objectives to be considered in concurrency control protocol in mobile database environment.

III. Proposed Method

Some concurrency control scheme performs better in some particular database systems e.g. Time stamp based multi-version protocols is better in real time databases. So according to the demand of the mobile database, we can define some specific method for a defined type of transactions and consider its condition to set best performed rule for it. We can categorize transaction operations as:

Transactions that only Read an item (TR), Transactions that only Write an item (TW), Transactions that Read as well as write an item that is Update (TU), Transactions that inserts an item (TI), Transactions that delete an item (TD) etc.

In my proposed method, I have used Time-stamp based multi-version protocol for TR and TW. So TR always gets the most recent version of the data elements and TW just creates a new version of data element and the old version remains intact and hence possibility of occurring conflict is zero.

For the TU, writing conflict may happen. So it is essential to resolve the conflicts among the TUs effectively to maintain integrity of the data item.

In my proposed method, when a client MH requests for the data items e.g. Xj. On which the MH going to have write operation(s), Server stores an entry for that MH including the client Id, Time of the beginning of transaction execution (TB), the list of data items Xi for that client, maximum validation period PV, Rank RC.

T_B is maximum duration to execute and send cache to server which is supplied by MH for the use of server to achieve more concurrency. PV is a time limit within which MH have to perform its operation and send cache to server to commit else request will be discarded from the transaction list of the server. PV is determined by the TB and the bandwidth of the connection to that client. RC is an integer that keeps track of the commit attempt of a transaction.

As all transactions are executed locally affecting the cache of the client, client sends partial update of the transaction for a data item that has no further update operation within this transaction for the partial commitment to the server if connection between mobile client and server be available. It would increase the concurrency significantly. In this case server checks for the conflicts and resolves by commitment algorithm. If connection is not available during execution then at the end of all updates in client cache, updated cache is sent to server to be committed. All kind of update is done according to the commit algorithm described below:

When any client MH_i sends cache to server, server finds the clients MH_a, MH_b with which MHi conflicts. Server performs following operations:

If MH_i have expired maximum validation period of MH_i, Server cancels the commit operation and causes to restart the transaction of MH_i.

If Executed Time Tex of $MH_i < \sum j$ Tex of MH_j, for all j X_i
Then
If RC $> \sum j$ RC for all j Xi
MH_i commits; invalidating the conflicting clients causing to restart their execution immediately with updated value of the data item.
Else
MH_i aborts and restarts its execution.
Else
MH_i commits and causes to restart their execution immediately with updated value of the data item.

IV. Performance Analysis

The proposed concurrency control scheme is assessed in this section. To perform assessment, it is compared with optimistic concurrency control methods as proposed technique is based on optimistic methods. In order to assess and compare performance of proposed technique, optimistic technique is simulated as well. I have used of network simulator NS-3 [22, 23] version 3.12 to implement my protocols.

Simulation is done using C + + in application layer protocols.

A circular-shaped area (D=1000 meters) space, a fix host (server), 5 mobile base stations, and 100 mobile clients are used in simulation environment. Mobile hosts can move randomly in various directions and communicate their requirements by mobile base station to the server.

Simulation is done for getting the comparison result values for the optimistic method and proposed scheme. Since response time and waiting time of a transaction are two important parameters of performance, the graph is drawn and studied only for the two parameters. The load of the system is increased incrementally.

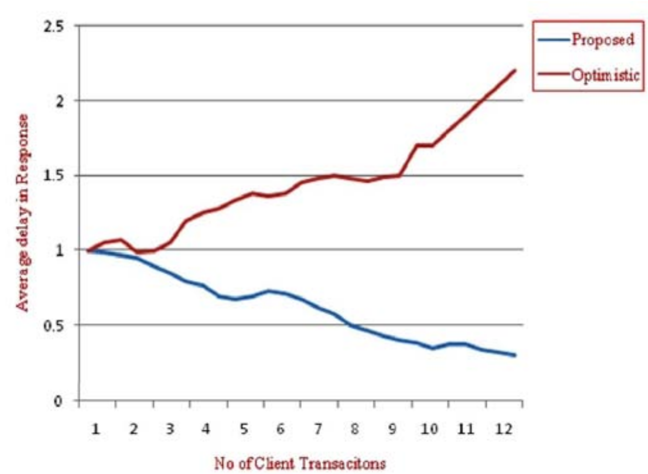


Figure 2 : Comparison of avg delay in response with number of client transactions

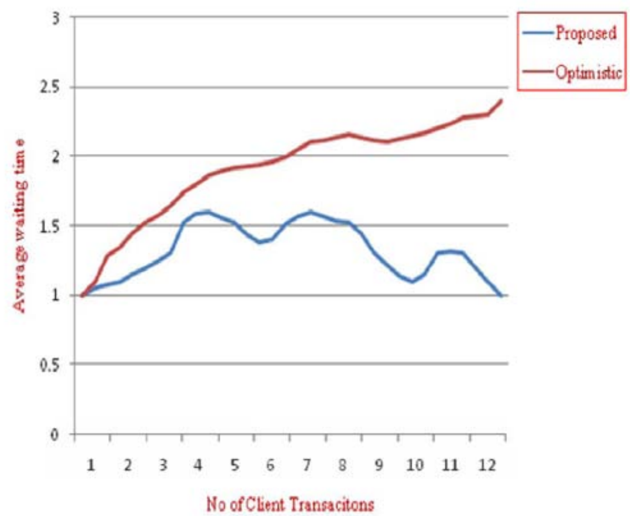


Figure 3 : comparison of avg waiting time with number of client transactions

In the figures Fig. 2 and Fig. 3 the comparison of the average delay in response as well as the avg waiting time with the increase in number of client transactions is shown. Average delay in response time is the average increase in normal transaction execution time due to concurrency. Waiting time of the transaction is the time for which temporary hiatus in execution of the transaction appear due to non-availability of shared data items. With the increase of the number of transactions avg. delay in response and waiting time increases significantly whereas proposed method performs better.

V. Conclusions

In this paper, a new concurrency control technique based on optimistic method in mobile database environment with client-server model is introduced. This method is based on Optimistic method, so there is no blocking or deadlock in transactions. Moreover, concurrency Degree of transactions is high. To reduce the rate of abortion, proposed algorithm, based on broadcast commit, applies the technique of partial update as well as assigning priority depending on some parameters. This leads to increase in the rate of commitment of transactions. There some overhead for the determining execution time of transaction which is worthless and is tolerable. There is plan to improve this method by introducing time stamp in it.

References Références Referencias

1. Joe, C.H.Y., Chan, E., Lam, K.Y., Leung, H.W, "An Adaptive AVI-based Cache Invalidation Scheme for Mobile Computing Systems", DEXA 2000, pp.155-178, IEEE Computer Society, USA.
2. Ho-Jin Choi, Byeong-Soo Jeong, "A Timestamp-Based Optimistic Concurrency Control for Handling Mobile Transactions", ICCSA 2006, LNCS 3981, pp. 796 -805, 2006.
3. K. Ali and b. Ahmad," A Concurrency Control Method Based on Commitment Ordering in Mobile Databases", International Journal of Database Management Systems (IJDMS) vol.3, no.4, November 2011.
4. Minsoo Lee, Sumi Helal, "HiCoMo: High Commit Mobile Transactions", Distributed and Parallel Databases, Vol.11, pp.73 -92, 2002, Kluwer Academic Publishers.
5. Mohammed Khaja Nizamuddin, Syed Abdul Sattar, "Adaptive Valid Period Based Concurrency Control In: Recent Without Locking in Mobile Environments" Trends in Networks and Communications, Springer CCIS, Vol.90, Part 2, pp 349-358, 2010, Springer Heidelberg.
6. Nadia Nouali, Anne Doucet, Habiba Drias, "A Two-Phase Commit Protocol for Mobile Wireless the Australasian Database Environment", Vol. 39, 16 Conference, 2005.
7. Patricia Serrano-Alvarado, Claudia Roncancio, Michel Adiba, "A Survey of Mobile Transactions", Distributed and Parallel databases, 16,193-230, 2004, Kluwer Academic Publishers.
8. Salman Abdul Moiz, Dr. Lakshmi Rajamani, "An Algorithmic approach for achieving Concurrency in Mobile Environment", INDIACom, 209-211, 2007.
9. K Vijay, "TCOT -A Timeout-Based Mobile Transaction Commitment Protocol," IEEE Transactions on Computers, vol. 51, pp. 1212-1218, 2002.
10. S.Moiz and L. Rajamani, "Single Lock Manager Approach for Achieving Concurrency Control in Mobile Environments," in 14th International Conference on High-Performance Computing, Goa, India, 2007, pp. 650-660.
11. M. Salman Abdul, "Concurrency Control Strategy to Reduce Frequent Rollbacks in Mobile Environments," in International Conference on Computational Science and Engineering, 2009, pp. 709-714.
12. S. A. Moiz and M. K. Nizamuddin, "Concurrency Control without Locking in Mobile Environments," in Emerging Trends in Engineering and Technology, 2008. ICETET '08. First International Conference on, 2008, pp. 1336-1339.
13. D. L. R. Salman Abdul Moiz, "A Real Time Optimistic Strategy to achieve Concurrency control in Mobile Environments using on-demand multicasting," International Journal of Wireless & Mobile Networks (IJWMN), vol. 2, pp. 172-185, 2010.
14. S. Madria, M. Baseer, V. Kumar, and S. Bhowmick, "A transaction model and multiversion concurrency control for mobile database systems," Distributed

and Parallel Databases, vol. 22, pp. 165-196, 2007.
15. S. K. Madria, M. Baseer, and S. S. Bhowmick, "A Multi-version Transaction Model to Improve Data Availability in Mobile Computing," in International Conferences DOA, CoopIS and ODBASE 2002, pp. 322-338.
16. C. Sung Ho, L. JongMin, H. Chong-Sun, and L. WonGyu, "Hybrid concurrency control for mobile computing," in High Performance Computing on the Information Superhighway, 1997. HPC Asia '97, 1997, pp. 478-483.
17. A. C. Carmine Cesarano, Vincenzo Moscato, Antonio Picariello, Antonio d'Acierno, "A Hybrid Approach for Improving Concurrency of Frequently Disconnecting Transactions," International Journal of Computer Science and Network Security (IJCSNS), vol. 7, pp. 205-215, 2007.
18. A. Chianese, A. d'Acierno, V. Moscato, and A. Picariello, "Pre-serialization of long running transactions to improve concurrency in mobile environments," in Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on, 2008, pp. 129-136.
19. C. S. Lee, K.-W. Lam, and S. H. Son, "Concurrency Control Using Timestamp Ordering in Broadcast Environments," The Computer Journal, vol. 45, pp. 410-42 2, January 2002.
20. C. S. Lee, K. Wa Lam, and T.-W. Kuo, "Efficient validation of mobile transactions in wireless environments," Journal of Systems and Software, vol. 69, pp. 183-193, 2004.
21. Salman Abdul Moiz ,Dr. Lakshmi Rajamani, "Concurrency Control Strategy to Reduce Frequent Rollbacks in Mobile Environments," CSE, vol.2, pp.709-714, 2009 International Conference on Computational Science and Engineering, 2009.

Global Journals Inc. (US) Guidelines Handbook 2013

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

- 'FARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FARSC" can be added to name in the following manner. eg. **Dr. John E. Hall, Ph.D., FARSC or William Walldroff Ph. D., M.S., FARSC**
- Being FARSC is a respectful honor. It authenticates your research activities. After becoming FARSC, you can use 'FARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 60% Discount will be provided to FARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- FARSC will be given a renowned, secure, free professional email address with 100 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- FARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 15% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- Eg. If we had taken 420 USD from author, we can send 63 USD to your account.
- FARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- After you are FARSC. You can send us scanned copy of all of your documents. We will verify, grade and certify them within a month. It will be based on your academic records, quality of research papers published by you, and 50 more criteria. This is beneficial for your job interviews as recruiting organization need not just rely on you for authenticity and your unknown qualities, you would have authentic ranks of all of your documents. Our scale is unique worldwide.
- FARSC member can proceed to get benefits of free research podcasting in Global Research Radio with their research documents, slides and online movies.
- After your publication anywhere in the world, you can upload you research paper with your recorded voice or you can use our professional RJs to record your paper their voice. We can also stream your conference videos and display your slides online.
- FARSC will be eligible for free application of Standardization of their Researches by Open Scientific Standards. Standardization is next step and level after publishing in a journal. A team of research and professional will work with you to take your research to its next level, which is worldwide open standardization.

- FARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), FARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 80% of its earning by Global Journals Inc. (US) will be transferred to FARSC member's bank account after certain threshold balance. There is no time limit for collection. FARSC member can decide its price and we can help in decision.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

- 'MARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'MARSC" can be added to name in the following manner. eg. Dr. John E. Hall, Ph.D., MARSC or William Walldroff Ph. D., M.S., MARSC
- Being MARSC is a respectful honor. It authenticates your research activities. After becoming MARSC, you can use 'MARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 40% Discount will be provided to MARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- MARSC will be given a renowned, secure, free professional email address with 30 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- MARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 10% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- MARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- MARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), MARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 40% of its earning by Global Journals Inc. (US) will be transferred to MARSC member's bank account after certain threshold balance. There is no time limit for collection. MARSC member can decide its price and we can help in decision.

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJCST for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

MANUSCRIPT STYLE INSTRUCTION **(Must be strictly followed)**

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, “Abstract” word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.
Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript’s Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications.

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also.Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

You must strictly follow above Author Guidelines before submitting your paper or else we will not at all be responsible for any corrections in future in any of the way.

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

1. Choosing the topic: In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. **Pick a good study spot:** To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. **Think technically:** Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page
- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.

- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.
- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.
- Do not give permission to anyone else to "PROOFREAD" your manuscript.
- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

INDEX

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350