Certain Investigations on Query Processing Related Issues in Mobile Computing Platforms for Large Datasets

T.P.Andamuthu

Professor Department of ECE, Maharaja Engg College For Women, Perundurai, Tamil Nadu.

Dr.P.Balasubramanie

Professor, Department of CSE, Kongu Engg. College, Perundurai, Tamil Nadu

Abstract

Applications in real time environment demands interactions from different devices both fixed and mobile devices in a seamless manner. The users interacting are in different geographic environments. The datasets are of different nature and allows for interaction across users in different levels with more levels of constraints. This in turn leads to different directions for addressing the issues in mobile computing. The queries from users are also complex and involve joins. The data residing also leads to issues of data management and prioritization of requests in communication. The technology enabled devices are also improving with processors and different platforms like HTML5 to address the existing issues. This paper presents a detailed survey of different issues in query processing in mobile environment. The solutions are also presented for the existing issues at the end of the paper.

Keywords: Dataset, platform, join, query, optimization

1. Introduction

Mobile computing plays a vital role in today's communication. The participants in mobile computing include mobile devices, mobile users and the mobile providers. The data that needs to be processed resides in the storage device. The requests are manipulated in form of queries. These queries pass across different nodal devices that carry the data. The relational algebra operations are applied on the database for presenting the results. The way of handling the queries depends on the type of database considered. The data requested can be a single type or multimedia content residing in the server.

2. Existing Work

Margaret H.Dunham analyzed the approaches that exist for mobile computing. The conclusion was that the approach for mobile computing is similar to the approaches in distributed computing and systems. Three different applications were analyzed and the Solutions were suggested for handling data management issues namely Applications, Transactions, Recovery Replication, Query processing and Name

Resolution. Distributed Transactions were not addressed in this work. JM Hellerstein analyzed the approaches for adaptive query processing. The existing approaches were compared for frequency adaptability. Adaptive Query operators were suggested for handling block operations. The analysis gave suggestions for designing appropriate data structures for query processing.

Tolga Urhan suggested the approach of using XJoin for applying pipelined join operations. The suggestion was that the XJoin could resolve the conflicts in handling multiple queries in applications. Vinayak Borkar proposed ALDSP architecture for handling mobile transactions. The approach was designed for handling credit card information in mobile data. The query processing was done using parsing, expression tree construction, normalization, type checking, Optimization, code generation and Query execution phases. XQuery Source was used for analysis. Haifeng Jiang suggested the performance estimate of twig queries for mobile systems. The performance estimation was done based on number of scanned elements, number of disk I/O,CPU time. LRU Buffer manager has been used for buffer management. The approach could be extended and implemented in simulators like ns2 and the performance could be estimated.

Mihaela A.Bornea proposed a Diner algorithm for handling join operations in Queries. The proposed Diner algorithm has been proved to be efficient in terms of handling burst data, memory limitations, and faster output rate. The approach could be implemented for real time applications and studied. Tarkoma.S analyzed the importance of different mobile computing platforms. The study gave importance and directions for users to developing applications through HTML5. HTML5 combines design, user interface development and data management in a single platform. HTML5 also supports for managing requests efficiently in different mobile computing devices and applications. Ahmed Alghamdi, suggested the concept of using Store and Forward protocols for Global Positioning System for evaluating the throughput for a mobile adhoc network. The simulations were done using ns2.

The Table 1 below presents the comparative approaches for query processing in mobile environment.

Table 1: Comparative approaches for Query processing in Mobile Environment

Author ,Title and Year	Approach/Methodology used	Merits	Limitations	Application	Performance parameters	Comments
Anurag Kahol et al (2001)	Data structure using time cache, first_request, first_waiting , status of is being designed	Modeling of invalidation messages. Reduces uplink request and average latency. State information is	Mobility patterns are not considered for Hit Ratio, latency, cost of invalidations.	General datasets	Hit ratio, Mean Query Delay	1. Overhead is reduced in managing the cache.

	Cache Maintenance Scheme is proposed	maintained in Home MSS. Data loss is reduced. Bandwidth saving and data access latency.				
Pauray S.M.Tsai(2002)	Response Time model for estimating the cost of the queries	Reduction in Response Time. Foreign functions for select queries is considered. Works for the semi-join approach.	Do not work for foreign joins Parallel processing is not considered	General datasets	Join selectivity, Percentage	Could be extended for cluster machines
Henrique Andrade(2004)	Database Optimization framework.	Improves Computation reuse, Query scheduling, semantic caching. Analyzed for shared memory multiprocessor, workstation cluster, Grid Environment Proxy for handling cached aggregate and automatic sub queries.	Query Execution plan, Compiler Optimization is not handled.	General datasets	Batch Execution Time	Could be applied for real-time datasets
Yongluan Zhou(2005)	Dynamic Cost Model	A new cost model is proposed. A new metric named performance ratio is suggested. Minimization of communication cost.	Large Complex queries are not considered. Batch analysis is not considered.	General datasets	Performance Ratio Adaptability Stability Sensitivity	Foreign Functions could be applied. Complex datasets could be considered.
Wen-Chin Peng(2005)	A Systematic procedure for MI/SI semijoins.	Asymmetric features for computing capability between server and mobile computing, sending and receiving, activeness & idleness are considered.	Semi joins is not considered Average complexity is not measured	General datasets	Processing ratio, Energy consumption, Idling coefficient, Processing time, Data transmission, energy consumption	Could be extended to complex joins with foreign functions
Stratos Idreos et al (2006)	System Model based on CHORD DHT protocol and Double Attribute based Index algorithms.	Considers continuous two way equijoin queries. An appropriate algorithm has been suggested for Indexing. Network Traffic is reduced.	Foreign Functions has not been considered Scalability factor has not been addressed. Not addressed for Internet Specific applications.	General Dataset	Network Traffic Traffic Cost Bos Ratio Load Distribution	Could be implemented for P2P Databases, Locality aware DHT's
Songting Chen(2006)	Dependency Detection and correction algorithms	A Novel framework named DyDa is proposed.	Three types of data anomalies namely data updates, data preserving schema changes, non data preserving schema changes is		Abort Cost, Varying interval	Performance ratio can be estimated.

			considered. Dynamic Data and Scheme management Minimal Overhead			
Bugra Gedik(2006)	MobiEyes- A Distributed real time Location monitoring System	Mobile Objects is also considered . High degree of precision . Concrete data structures, Distributed coordination mechanism	Has not been considered for real time datasets.		Messaging Cost, server load, Amount of computation	Real time datasets could be considered.
Benjamin Arai(2007)	Sampling Based Techniques for answering adhoc queries Two phase adaptive sampling approach for graph	Count aggregations are applied for aggregate functions.	Aggregation of various topologies Considers Jump handling for resolving conflicts		Accuracy Clustering level	Can be tested for real time datasets
Manish Sethi(2008)	Open Framework for handling Metadata	Web service for handling data management	Servlet for Query management Suggested the analysis for J2EE app server for web application. Mapping of the model to real dataset analysis is missing.			Could be extended for real time datasets for analysis.
Jian Li,Amol et al(2011)	A Novel Algorithm for analyzing the data movement	Predictions like – solving optimization ,problems in Polynomial time were considered.	Response time is improved	Dataset of sensor Network and random datasets	Cost Estimation of clusters and Queries	Could be applied for real world data
Lijiang Chen et al(2011)	A Novel algorithm named PaDaSkyline for parallel Query Processing	Three Methods namely MaxSum,MaxDist,IPV MaxSUm,IPVMaxDist are compared.	Query Traffic, Response Time, Data Reduction Rate is improved.		Cost Model, Response Time, Precision, Query Traffic, Data Reduction Efficiency Distance Threshold	Could be applied for real world datasets
Chi-Yin Chow et al (2011)	Framework(ContP) for managing Quality issues in Mobile Adhoc Networks	Qos profile Generation User can provide the quality requirement Scalable framework for supporting more number of users. Qos Measurement s for the local cache Maintains Query Answers.	Communicatio n Overhead is increased	Grid based Data sets	Power consumption, message transmission level	Could be tested for data transmission on real requests in Mobiles
Andre Freitas(2012)	Query and Search Solution for Linked Data	Solutions for Query handling and vocabulary issues are addressed.	Has not considered for real time implementatio n- only		Analyzed the challenges for Entity centric search, Question answering,	Could be extended and integrated into applications requiring

www.ijcait.com

International Journal of Computer Applications & Information Technology
Vol. I, Issue III, November 2012 (ISSN: 2278-7720)

			scenario analysis is performed		Structure search, best effort natural language.	learning.
Ricardo J.Barrientos(2012)	MultiGPU Metric Spaced Technique	The proposed approach outperforms Open MP and GPU based platforms.	Has not considered real-time issues in network communicatio n.	Images	List of Clusters, SSS-Index	Could be applied to any domain specific problems
Sergio Illari et al (2006)	System for processing Location dependent queries	General solution for processing queries . Efficient Processing with optimization support. Tracking of objects Decentralized solution for data monitoring. Mobile Agent is used.	Location dependent constraints are not analyzed.	Data sets concerned to car.	Scalability Query Size Location Error	Could be applied for real data sets gathered from mobile communication
Cecilia Mascola(20 01)	XMIDDLE a peer to peer middleware is being designed for replica maintenance	Transparent sharing among heterogeneous peers. XMIDDLE analysis for replica management.	Does not consider for any real-time request processing.	No datasets is considered		Could be analyzed for any real time application.
Reza'B'Far(2005)	Mobile Agent based approach for mobile computing	Usage of UML for Data replication and Synchronization issues. Mobile Agents for computing.	Does not consider for any applications.			Could be extended and applied for learning the performance in specific applications.
Karin Coninx(200 3)	Dygimes Framework	WebService Description Language for User Interface Designs. Supports for Remote Method Invocation. Customization of Layouts.	Has not handled for any specific widget			User centric request and operation for widgets could be addressed.
A.Vakali(20 05)	XML Datastores	Scalability, Data-Access Speed and reliability. Analysis of different XML stores. Implicit data caching.	Has not be analyzed for different requests received in a mobile environment			Relations could be defined and the performance could be measured.
G.H.Forman (1994)		Analysis on Data. Dissemination analysis, location dependent Querying, Advanced interfaces for computers.	Has not addressed solutions for prototyping, Bandwidth Utilization, Transactional properties, Optimization of location dependent query, Data visualation			Could be extended for language specific platforms and addressed.

4. Outcomes of the Survey

The survey has presented the following observations:

- Mobile computing applications exist and have the same behavior as distributed applications.

- Applications developed in latest platforms like HTML 5 can support for multiple request issued from different mobiles.
- The XQuery can work with XML more efficiently for handling the mobile databases. Servlet can also be integrated to the applications.
- Buffer Management is inherently handled using Buffer Managers.
- Algorithms developed for joins could be implemented and analyzed for real time data sets.
- Mobile simulators like ns2 could also be used such the query processing is handled using simulations and the traffic could then be learnt.
- There needs to be a balance to be achieved between network request, bandwidth and query requests issued in the network.
- Network Failure is inherently handled by the language platform like HTML5 and the request is reissued for retrieving the data from appropriate databases.
- The complexity of the queries depends on the domain of the application and the type of request issued by the users.
- There could be a more better way of managing the mobile request by mapping XML to HTML and reverse mapping in a object oriented manner.
- Caching is inherently available in the mobile devices for storing and managing the data.

5. Conclusion

This paper has presented a detailed survey on the query processing related issues in different directions like databases, applications, query complexity, mobile platforms, and mobile centric language platforms. The survey concludes that the users developing and writing mobile applications could choose an appropriate environment and build on queries for handling the data requests in an object oriented manner. Appropriate algorithms could be selected for handling the join related issues and data management issues. A whole system for a specific application could be developed and tested with the best practices suggested as a part of the survey.

6. References

1. Anurag kahol,Sumit Khurana,Sandeep K.S. Gupta, Pradip K.Srimani ,"A strategy to manage Cache consistency in a Disconnected Distributed Environment" ,IEEE Transactions on Parallel and Distributed Systems, Vol.12,No.7,July 2001.
2. Pauray S.M.Tsai,Arbee L.P.Chen " Optimizing Queries with Foreign Functions in a Distributed Environment" , IEEE Transactions on Knowledge and Data Engineering,Vol 14,No.4,July/August 2002.
3. Henrique Andrade ,Tahsin Kurc,Alan Sussman,Joel Saltz ,"Optimizing the Execution of Multiple Data Analysis Queries on Parallel and Distributed Environments",IEEE Transactions on Parallel and Distributed Systems,Vol.15,No.6,June 2004.
4. Yongluan Zhou,Beng Chin Ooi, Kian-Lee Tan, "Dynamic Load Management for Distributed Continous Query Systems", proceedings for the 21st International Conference on Data Engineering(ICDE'2005), 2005
5. Wen-Chin Peng,Ming-Syan Chen " Query Processing in a Mobile Computing Environment: Exploiting the Features of Asymmetry" , IEEE Transactions on Knowledge and Data Engineering, Vol.17,No.7,July 2005.
6. Stratos Idreos,Christos Tryfonopoulos,Manolis Koubarakis,"Distributed Evaluation of Continous Equi Join Queries over Structured Overlay Networks", Proceedings of the 22nd IEEE International Conference on Data Engineering,2006.
7. Songting Chen,Xin Zhang,Elke A.Runensteiner ,"A Compensation-based Approach for view Maintenance in Distributed Environments", IEEE Transactions on Knowledge and Data Engineering, Vol.18,No.8,August 2006.
8. Bugra Gedik,Ling Liu , "MobiEyes :A Distributed Location Monitoring Service Using Moving Location Queries" , IEEE Transactions On Mobile Computing , Vol.5,No.10,October 2006.
9. Benjamin Arai,Gautam Das,Dimitros Gunapulos,Vano Kalogarekai,"Efficient Approximate Query processing in Peer to Peer Networks",IEEE Transactions on Knowledge and Data Engineering,Vol.19,No.7,July 2007.
10. Manish Sethi,Ashok Anand,Dipayan Gangopadhayay,Venkateswara Reddy,Manish Gupta,"An Open Framework for Federating Integrated Management Model of Distributed IT Environment" ,IEEE 2008.
11. Jian Li,Amol Desphande,Samir Khuller,"Minimizing Communication Cost in Distributed Multi Query Processing" ,IEEE International Conference on Data Engineering,2011.
12. Lijiang Chen,Bin Cui,Hua Lu,"Constrained Skyline Query Processing against Distributed Datasites", IEEE Transactions on Knowledge and Data Engineering, Vol.23,No.2,Febraury 2011.
13. Chi-Yin Chow,F.Mokbel,Hona Va Leong,"On Efficient and Scalable Support of Continuous Support of Continous Queries in Mobile Peer to Peer Environments" ,IEEE Transactions on Mobile Computing,Vol.10,No.10,October 2011.
14. Andre Freitas,Edward Curry,Joao Gabriel Oliveira,Sean O Riain, "Querying Heterogenous Datasets on the Linked Dataweb- Challenges Approaches and Trends",IEEE Internet Computing,2012.
15. Ricardo J.Barrientos ,Jose I.Gomez,Mauricio Marin,"Range Query Processing in a multi-GPU Environment" ,Proceeding of the 10th IEEE International Symposium on Parallel and Distributed Processing with Applications", 2012
16. Sergio Illari,Eduardo Mena, Aranta Illarramendi,"Location-Dependent Queries in Mobile Contexts : Distributed Processing Using Mobile Agents",IEEE Transactions on Mobile Computing ,Vol.5,No.8,August 2006.
17. Cecilia Mascola,Licia Capra,Wolfgang Emmerich,"An XML based Middleware for Peer to peer computing" Proceedings of the First international Conference on Mobile Computing" ,2001.

www.ijcait.com

International Journal of Computer Applications & Information Technology
Vol. I, Issue III, November 2012 (ISSN: 2278-7720)

18. Reza'B'Far,"Mobile Computing Principles – Designing and Developing Mobile Applications with UML and XML",Cambridge University press,2005.

19. Karin Coninx,Kris Luyten,Chris Vandervelpan,Jan Van den Bergh,bert Creemers,"Dygimes:Dynamically Generating Interfaces for Mobile Computing Devices and Embedded Systems", Lecture Notes in Computer Science, 2003, Volume 2795/2003, PP. 256-270.

20. A.Vakali,Catania.B,Maddalena,"XML Data Stores:emerging practices ",IEEE Conference on Internet Computing,2005.

21. G.H.Forman,John Zahorjan,"Challenges of Mobile Computing" IEEE 1994.

22. Daniel Barbara, " Mobile Computing and Databases – A Survey", IEEE Transactions on Knowledge and Data Engineering,Vol.11,No.1,January/February 1999.

23. Vassillis Tsaoussidis,Ibrahim Matta,"Open Issues on TCP for mobile computing", Wireless Communications and Mobile Computing,2001

24. Margaret H.Dunham,Abdelsalam Helal,"Mobile computing and databases:anything new",AGM SIGMOD,Volume 24,Issue 4,Dec 1995.

25. JM Hellerstein,MJ.Franklin,S.Chandrasekararan,"Adaptive Query Processing : Technology in evolution", International Conference on Mobile Computing and Networking(MobiCom99).1999.

26. Tolga Urhan,Michael J .Franklin,"XJoin: A Reactively-Scheduled Pipelined Join Operator" ,Data Engineering,June 2000, PP:27-33

27. Vinayak Borkar,Michael Carey,Dmotry Lychagin,Till Westmann,Daniel Engovatov,Nicola Onose,"Query Processing in the aqualogic dataservices platform" Proceedings of the 23nd International Conference on Very Large Databases 2006.

28. Haifeng Jiang,Hongjun Lu,Wei Wang,"Efficient Processing of XML twig queries with OR predicates" proceedings of the ACM SIGMOD International Conference on management of Data,2004.

29. Mihaela A.Bornea, Vasilis Vassaolos,Yannis Kotidis,"Adaptive Join Operators for Result Rate Optimization of Streaming Inputs",IEEE Transactions on Knowledge and Data Engineering,Vol22, No.8,August 2010.

30. Tarkoma.S,Lagerspetz.E,"Arching over the Mobile Computing Chasm:Platforms and Runtimes", IEEE computer 2011.

31. Ahmed Alghamdi, Raid Alghamdi,John DeDourek,Przemyslaw Pochec,"Store and Forward Protocol Advantage in a M2ANET Network", Proceedings of the Fourth International Conference on Advances in Future Internet,2012