AN EFFICIENT XML DATA STREAMING METHOD IN MOBILE WIRELESS NETWORK

[1]U. Ravindran, [2]J. Sribhavani,

Abstract – **XML documents are arriving at a rapid rate and the server is engaged for managing these documents and disseminating them to a pool of clients. An energy and latency efficient of XML dissemination scheme for the mobile computing is considered here. Therefore define a innovative unit structure called G-node for streaming the XML data in wireless environment. The system propose a weightless and sufficient cipher scheme, called Lineage Encoding, to guide the valuation of predicates and twig pattern inquiry over the stream. The global query processing time must also be reduced to provide fast response to the users. The XML broadcasting is done efficiently in such a way that the Server can support dynamic dissemination of a G node without any interruption in Broadcasting. The benefits of the structure indexing and attribute summarization that integrate relevant XML elements into a group and contribute a way for careful access of their quality values in a dynamic way that broadcasting can be done dynamically supporting Twig Pattern Queries.**

Index Terms -- **Twig pattern matching, wireless broadcasting environment, XML data streaming, parsing, XML Dissemination.**

I. INTRODUCTION

The accelerated improvement of cellular chain technologies, cellular mobile computing has become popular. Wireless telecasting is an sufficient instruction dissemination approach in the cellular mobile environment because of the following benefits [7], [9].

- *U.Ravindran,M.Tech., A.P/CSE Dept, ph-9840096489.*
- *SriBhavani.J is currently pursuing masters degree program in computer science engineering. Ph.9952505842.*

1) Server can device the huge no of client 2) telecasting channel is shared by huge client 3) client can accept information without posting call messages [23], [7]. XML scripts can be observed as arranged tree structure and then planned for data transaction. Cellular telecasting is an attractive method of disseminated data in mobile environment due to profitable characteristic. This cellular information telecasting supports full-text search operation by the mobile client [7],. The full-text search is one of the most popular query types used in various information system and access text retrieval. The important of XML data is more often inducing the needs for efficient process of XML. Twig pattern query commonly represented has a rooted, labelled XML query tree are used in the XML data [3]. Twig pattern relationship including parent-child (P-C) and ancestor-descendant (A-D) relationship.

The Extensible Markup Language (XML) has become a standard for data representation [2]. The ability to manage the huge collection of XML data and to discover knowledge from the web based information system. A potential result is to group the related XML data based on the tree structure. The results of an Xpath query are preferred by a position path. A position path consists of position steps. Handling each position step prefer a set of nodes in the documents tree that satisfy node and predicate the twig pattern query. In the conventional query processing, efficiency and scalability are major concerns in the wired environment, mobile client use battery powered mobile device and the energy conservation of mobile device.

The major goal of conventional query processing is streamed, XML data are to minimize the computation costs and filtering time. Focus of our work is how to reduce the downloads of data from the cellular channel. Broadcasting is the simultaneous transmission of the same message to multiple recipients [2], [18]. Telecasting arise when transferred data packet is accepted by all grid machine. Security issues may occur at the time of telecasting and lead to data damage if stormed by intruders. There are two common telecasting statuses for data telecast 1) Telecasting status- data is regularly telecast on the lower-link channel.

ISSN: 2278 – 1323

Clients only "observe" to that channel and down-load data they are excited. 2) On-Demand status-the clients send their calls to the server over upper link channel and the server examine all awaiting calls to determine the data of next broad-cast cycle.

The machinery proposes the energy and latency efficient XML propagate arrangements for the mobile computing. We define a innovative unit structure called G-node for streaming XML data in the cellular environment [7], [13]. It accomplishes the profit of the structure indexing and attributes summarization that can coordinate the related XML aspects into a cluster. It implements a action for careful access of their quality values and passage content. We also propose a weightless and sufficient cipher scheme, called Lineage Encoding, to guide the valuation of predicates and twig pattern query over the cascade. The Lineage Encoding scheme perform the parent-child relationships among XML elements as an arrangement of bit-strings called Lineage Code (V, H), and implements basic operators and functions for major twig pattern inquiry handling at mobile clients [7]. Pervasive experiments using absolute and fabricated data sets determine our device to perform the traditional cellular XML telecasting methods for transparent path query well as composite twig pattern query with acknowledge condition.

II. RELATED WORKS

The adequate, sufficient XML data telecasting method in a mobile cellular grid. It is slowly developed the demand of XML, sufficient method to transfer XML data over cellular telecasting environments is immediate prescribed. This machinery is a innovative XML data cascading method for cellular telecasting environments is proposed. An XML stream is organized to enable a discriminatory approach device for transparent XPath queries, by borrowing the path compact performance, which was initially, formulate for indexing semi-structured data. In order to handle framework knowledge as an XML stream index, the framework knowledge and data values of an XML document are partitioned [2], [12].

In observation decision, the machinery method demonstrates superior performance over previous approach with regard to both access and tuning time. The Advantage of this paper is 1)Path summary technique is used for accessing XML data in semi structured form 2) The structure information and text values of an XML document are separated and it is analysis for result. The Drawbacks of this document is 1)It proceeds long time for accessing data in xml semi structured form .2)It requires high bandwidth for getting appropriate result .3) Time and space complexity is more in semi structured for xml form. The title of Effective Scheduling Algorithm for On- Demand XML Data telecasts in cellular Environments. The configuration of data on cellular channels, which desire to diminish the connection time of mobile clients, is a key complication in data telecast. Many scheduling algorithms have been designed to organize at data on air. This machinery propose a innovative method to greatly reduce the tuning time by separated query results into XML fragment and to achieve better access adequate by combining related ones.

A. Prior Work

The data telecast scheduling problem of on-demand XML data broadcasts on the adequacy of a data item. Based on the solution, a Least Efficient Last (LEL) scheduling algorithm is also constructed to sufficiently organize XML data on cellular channels. Finally, we study the achievements of our algorithms over huge analysis. The decision show that our scheduling algorithms can diminish the both connection time and tuning time significantly when compared with existing work. The Advantage of this paper is Data broadcast is an efficient way for mutual knowledge transmission to a many number of mobile users. It offers great scalability, good power consumption, and efficient bandwidth usage.

The knowledge must allows an arbitrary number of clients to approach the data concurrently and appropriately applicable for heavily loaded systems Least Efficient Last (LEL) scheduling algorithm is also constructed to sufficiently formulate the XML data on wireless channels. The disadvantage of this paper is 1) The clients send their demand to the slave over higher-link channel and the server considers all pending requests to decide the contents of next telecast cycle.2) Telecast schedule determines what data items to be telecast by the slave and also the structure of data items on wireless channels so it takes more time to access.

B. Xpath and XQuery Language

XPath is a language for decision information in an XML document. It uses direction interpretation to preferred nodes in an XML document. The XPath language is based on a tree structure of the XML document, and implements the capacity to navigate around the tree. XPath expression specifies a pattern that selects a set of XML nodes [18], [16]. XQuery is described as the query and practical programming language that is constructed to query and convert the selection of analytical and non-analytical data, frequently in the form of XML [20], [21], text and with vendor

ISSN: 2278 – 1323

expansion for other data formats XQuery is a specification for a query language that allows a user or schedule to correct information.

C. XML Filtering

XML filtering structure is a goal to arrange quickly, on-the-fly coordinating of XML-encoded data to large numbers of query condition involve pressure on both structure and content [5]. A filtering system delivers the data to customer based on their interpretation activity. There are two fundamental sets of inputs to the system: customer profiles and document streams. Customer profiles define the information choice of individual users. These profiles may be constructed by the customer themselves, e.g., by selecting items in a Graphical User Interface, or may be constructed naturally by the system using machinery learning techniques. The other key inputs to a filtering system are the document streams simultaneously appearing the documents from data sources [8]. These documents are to be penetrated and delivered to customer or systems in a correct fashion. Filtering is performed by matching each arriving document against all of the user profiles to determine the set of interested users.

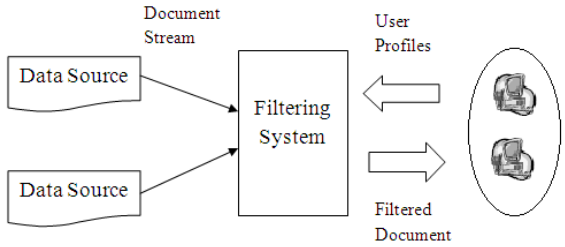

Fig 1: Overview or filtering system

D. XML Query Language

XML query language is a markup language for documents containing structured information [11]. Originally designed to meet the demanding of large-scale computerised broadcasting, XML is also performing a highly important role in the transaction of a wide variety of data. The increasing popularity of XML is partly due to the limitations of the other two technologies are Hypertext Markup Language (HTML) and standard generalized markup languages (SGML), for represents the structured and semi-structured documents.

HTML provides a fixed set of tags are mainly for presentation purpose and do not bear useful semantics while SGML is too difficult to implement for most application because of its complex specification. XML lies somewhere between HTML and SGML and is a simple yet flexible format derived from SGML. An XML identifiers enclosed in angle brackets. The most commonly used markup in XML data. Previous algorithm focus on XML tree pattern queries with only P-C and A-D relationship [7]. Little work has been done on XML tree queries which may contain wildcards, negation function, and order rule of all XML query. document always starts with a prolog markup. The minimal prolog contains a declaration that identifies the document as an XML documents. XML identifies data using tags, which are

The machinery calling an XML tree pattern with cancellation behaviour, wildcards and/or order as continued XML tree pattern. Four continued XML tree arrangement. Query (a) build a wildcard node which can equal any single hub in an XML database. Query (b) includes a negative edge. In query finds A that has a offspring B, but has no offspring C. In Xpath languages, the semantics of false edge can be granted with not Boolean behaviour. Query (c) has the regulated continued, which is equivalent to an Xpath //A/B following-siblings::C. The < in a box that all children under A are regulated. The accept of regulated-base tree arrangement is catch by a aligning from the arrangement nodes to nodes in an XML database such that the analytical and regulated relationships are fulfilled. Finally, query (d) is more complex, which contains wildcards, blank behaviour and arranged regulation.

```
1. <? xml version="1.0" encoding="UTF-8"?>
2. <bib>
3. <book>
4. <author> Suciu </author>
5. <author> Chen </author>
6. <title> Advanced Database System </title>
7. <chapter> <title> XML </title>
8. <section> <title> XML Specification </title>
9. <text> <keyword> markup </keyword> XML stands for…
10. </text>
11. </section>
12. </chapter>
13. </book>
14. </bib>
```

Fig 2: Sample XML Query Script

ISSN: 2278 – 1323

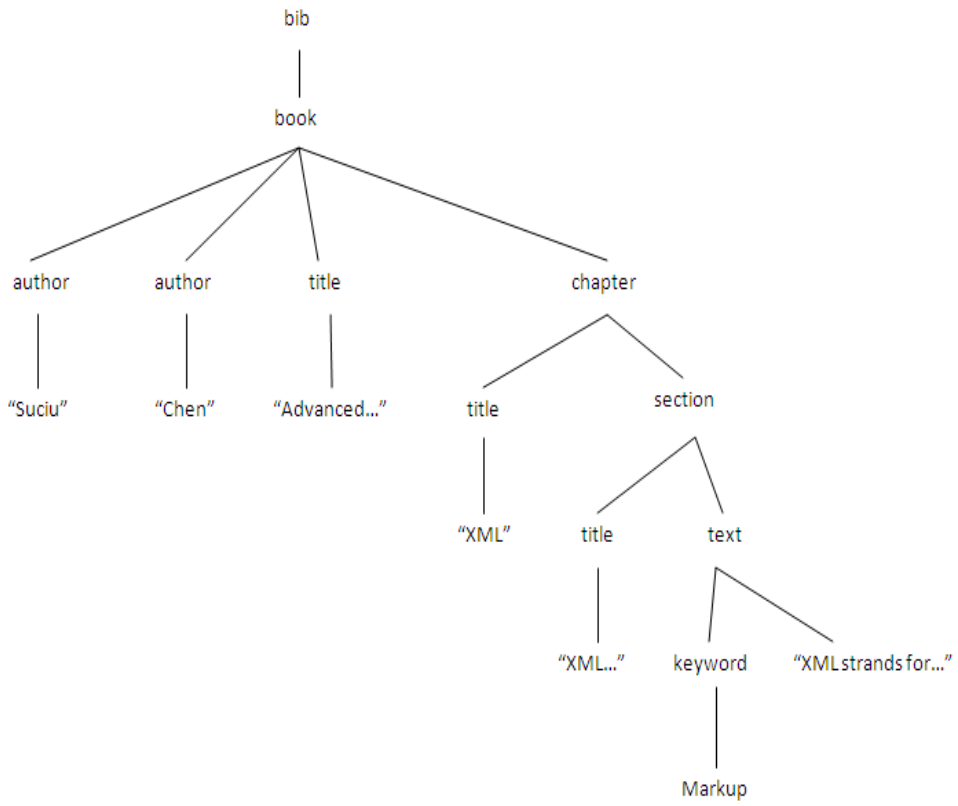

Fig 3: XML Query Tree Model

E. Twig Pattern Matching Algorithm

XML twig pattern identical is widely treated as a core activity in XML queries processing, there has been a rich set of twig pattern matching algorithm [11], [1]. Based on the containment labelling scheme, prior work decomposes a twig pattern into asset of binary exchange, which can be either parent-child or ancestor- descendant relationship. After that, each binary exchange is prepared using analytical join approach and final match results are achieved by combine single binary join results together. There is an increasing need for adequacy of handling the queries on XML data. Seeking for the occurrences of a tree arrangement query in an XML database is a core transaction in XML query handling. Prior works determine that holistic twig arrangement identical algorithm is an adequacy approach to report an XML tree arrangement with parent-child (P-C) and ancestor-descendant (A-D) exchange, as it can sufficiently authorise the size of transitional decision at the query handling. However, XML query expression decides high axes objective such as negation action, plan-based axis, and wildcards.

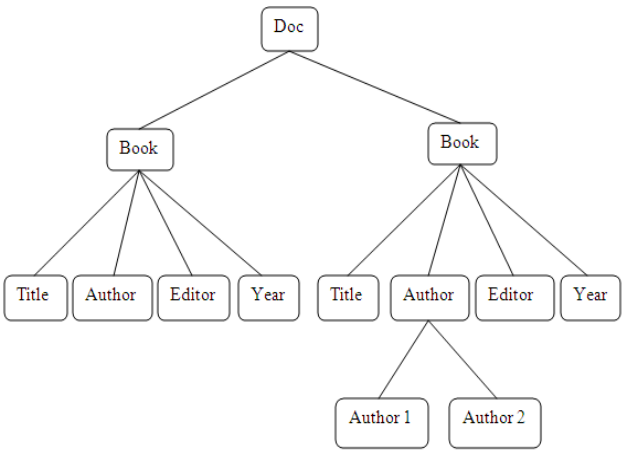

Fig 4: Tree Structure

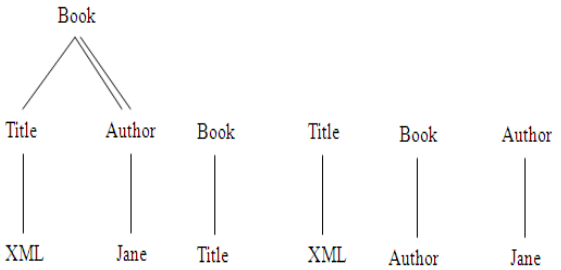

Fig 5: Structural Relationship

ISSN: 2278 – 1323

F. Holistic XML Twig Join

The sub-problem of answering twig queries efficiently, several algorithms based on the containment labelling scheme have been developed to process twig queries [4], [19]**.** Preceding work on XML twig arrange handling the dissolve a twig arrange into a set of binary exchange which can be either parent–child or ancestor-descendant relationship. After that each binary exchange is handled using analytical join capacity and the last event results are accessed by combine single binary join result together.

G. XML Documents Labelling Scheme

To designing a proper labeling scheme, the previous methods use a textual position of start and end tags or path expression. By applying these labeling schemes, one can determine the relationship e.g. ancestor-descendent and parent-child between two elements in XML documents from their labels alone introduce two most popular labeling schemes as follows. The importance of XML processing has become a significant field at present days with the intention to support user queries in the most proficient way. The elements in XML documents uniquely as well as preserve structural relationships among the nodes to cater queries with multiple combinations [2].

III PROPOSED WORK

Our work is different from the earlier ways, in this paper; we define a innovative unit structure called G-node for streaming XML data in the cellular environment [7], [17]. The machinery introduces a weightless and sufficient cipher scheme, called Lineage Encoding, to support the decision of predicates and twig arranged queries closed the stream. Figure 6 represent the architecture of wireless XML Broadcasting.

- First, the server can support a massive number of mobile clients without additional costs (i.e., scalability).
- Second, the broadcast channel is shared by many clients (i.e., - the effective utilization of bandwidth);
- Third, the portable clients can receive data without sending request messages that consumes much energy.
- Dynamic XML dissemination can be done so that the portable clients can have live updations of the data.
- No need to rely on third celebration achievements.

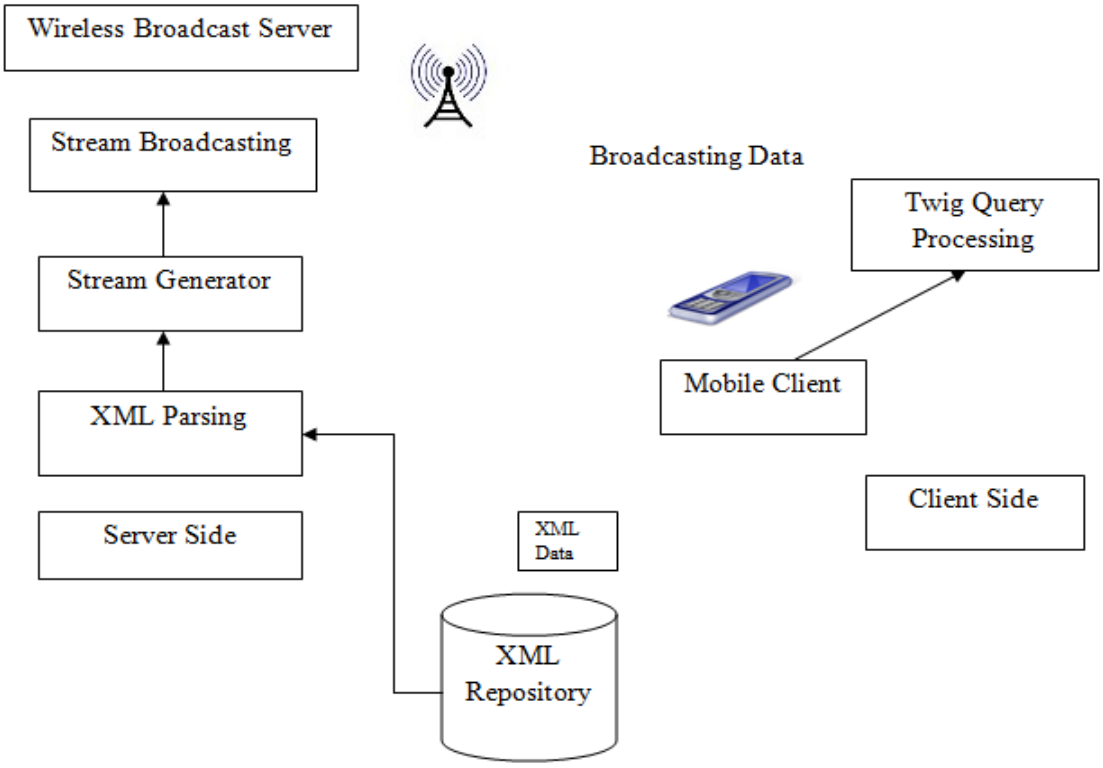

Fig 6: Architecture of wireless XML broadcasting

ISSN: 2278 – 1323

In cellular XML telecasting, the telecast server retrieves XML instruction to be disseminated from the XML repository. Then, it parses that instruction through parsing and generates a wireless XML stream. The XML stream is disseminated via a broadcast channel through structure indexing and attributes summarization [14]. In the client-side, if a query is issued by the mobile client, the mobile client tunes in to the telecast channel and selectively downloads the particular instruction from the XML stream for query processing. The XML telecasting is completed expeditiously in such the simplest way the server will support dynamic dissemination of a G node with none interruption in Broadcasting.

A. XML Data Manipulation

An XML data can be represented as a rooted, ordered, and labeled tree. Elements, quality, and texts are described by nodes, and the parent-child exchanges are described by edges in the XML tree. It shows a simple XML data that will be used as a running [6]. A server recover an XML data to be distributed from the XML repository and it develop a cellular XML steam by using SAX (Simple API for XML), which is an event-driven API. SAX invokes comfortable handlers during the parsing of an XML data. Analytical Indexing approach exchange many elements of the same path into one node, thus, the size of data stream can be decreased by redundant tag names thereby enabling Twig Pattern Query handling.

B. Innovative Encoding Scheme: Lineage Encoding

The innovative algorithm is divided into two main phases:

- Lineage Encoding, to device the queries associated with predicates and twig pattern coordinate (Fig 7). In the machinery introduce a two kinds of lineage codes, i.e., vertical code indicate by Lineage Code (V) and horizontal code indicate by Lineage Code (H), are used to perform parent-child links among XML aspect in two G-nodes [15]. It also decides a significant operators and functions that accomplish the bit-wise operations. Cellular XML streaming access that effectively supports twig pattern query handling in the cellular telecasting environment.

- The Attribute Value List (AVL) generated in Attribute Summarization with lineage encoded data is the key to process the Twig Pattern Queries in Selective tuning approach in the mobile end.

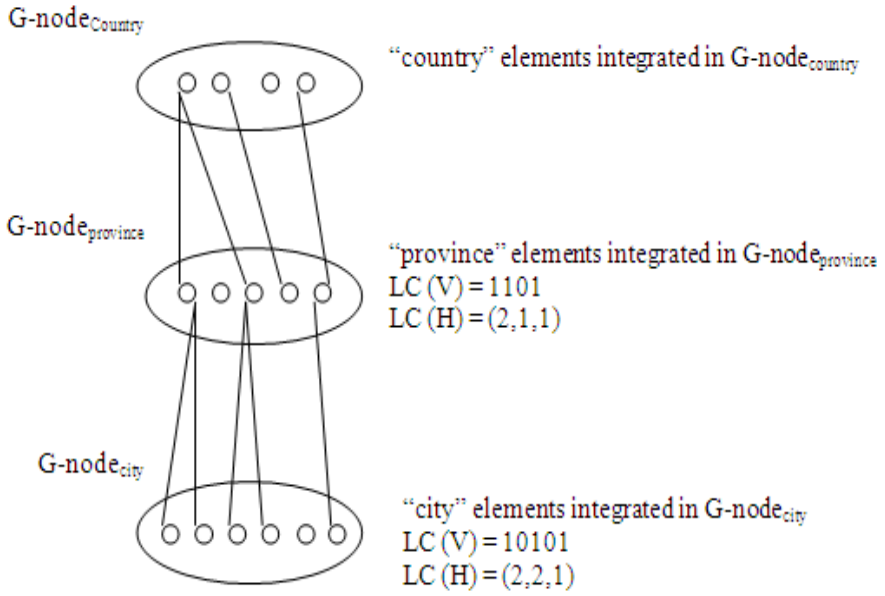

Fig 7: Lineage codes of G-node$_{country, G}$-node$_{province}$ and G-node$_{city}$

ISSN: 2278 – 1323

C. Creation of G-Node

The streaming unit of a wireless XML stream, called G-node. The G-node architecture defeats the complex hanging of XML data, and implements mobile clients to bound downloading of unwanted data during query handling. The group descriptor is a set of indices for careful access of a cellular XML stream. Node name is the tag name of combined elements, and section path is an XPath expression of combined elements from the starting node to the element node in the data tree. Child Index (CI) is a set of addresses that point to the starting positions of child G-nodes in the cellular XML stream. Attribute Index (AI) contains the set of attribute name and address to the root values of the attribute that are stored contiguously in Attribute Value list.

The components of the group descriptor are used to handle XML queries in the mobile client sufficiently. Especially, Node name and section path are used to identify G-nodes. Indices relating to time information such as CI, AI, and TI are used to download the next G-nodes, attribute values, and text. Finally, Lineage Code(V, H) is used to process the axis and predicate statement in the user's query & Attribute Value List store attribute values of the elements represented by the G-node, respectively .All the G-Node data's are telecasted with the help of a Wi-Fi appliances which can be accepted by any android appliances in its coverage.

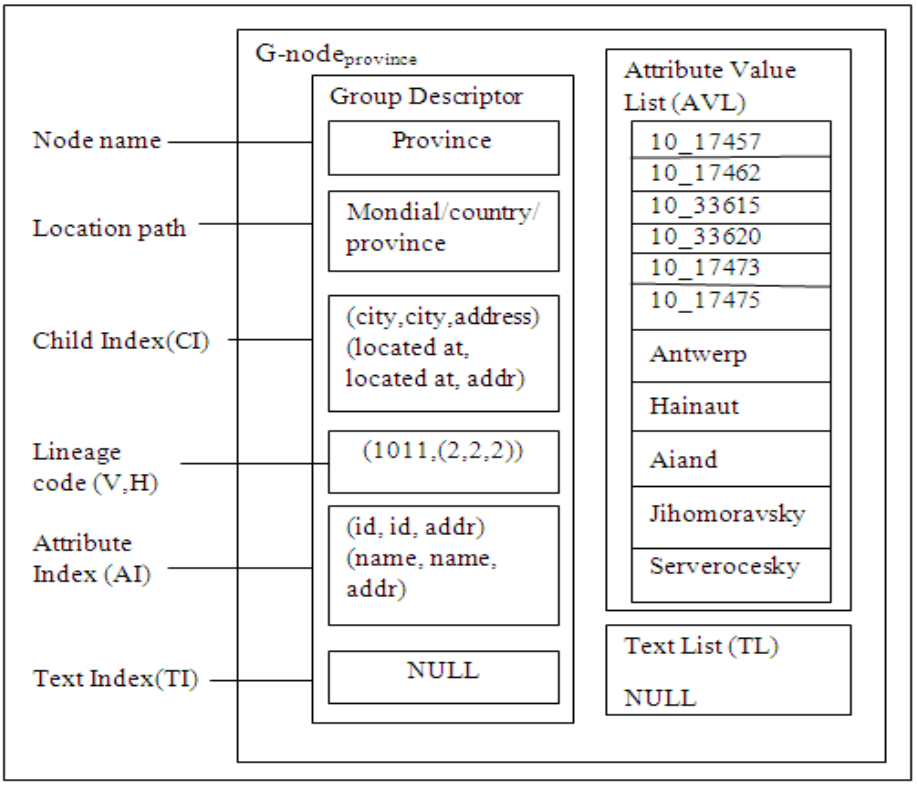

Fig 8: Structure of G-Node

D. Formation of Query Tree & Selective Tuning

In this category, the process reports how a mobile client can retrieve the document of its interests. Assign that there is no descendant axis in the customer query and query handling algorithms for an easy path query and a twig pattern query are given.

- Easy Path Query handling, Algorithm shows the easy path query handling over the cellular XML stream. Given a query, the mobile client to create a query tree. Then, it initializes to find applicable G-nodes over the cellular XML stream. The mobile client downloads a group descriptor of the G-node which corresponds to the query node. If the current node is the leaf node, the mobile client downloads.
- Twig Pattern Query handling the machinery cellular XML stream. In the Tree negation phase, the mobile client first

ISSN: 2278 – 1323

constructs a query tree [3]. Then, traversing the query tree in a depth-first manner, it selectively downloads group descriptors of the relevant G-nodes into the nodes in the query tree.

- Our Selective tuning approach is dynamic and it eases the client to minimize the tuning time and thereby reducing access time also. It dynamically chooses between the Twig Pattern Query and Normal Query and process to render the data.

- Tuning is optimized with the help of the XPath Query pattern which holds the predicates.

IV RESULT ANALYSIS

We present the enterprise results on the fabricated XML data set generated by XMark .It shows features of the streams generated by the considered methods. The disposed knowledge set has a huge number of component containing texts, whereas it has only a few quality. As a decision, the size of stream generated by LE is slightly larger than that of stream generated by PS. For enterprise, we used the XPath Mark query set, which is based on the XMark data set. Among queries in the XPath Mark query set, we omitted queries for testing namespace and node tests because these features are irrelevant to our work. Tested queries are separated by four types: easy path without predicates, easy path with predicates, twig pattern [11], ancestor-descendant axes. It shows all the queries used in our experiments.

It shows average access time evaluation results on the synthetic XML data set. For simple path without predicates queries, c-DIX shows the best performance. Since elements in c-DIX are ordered in an ascending order of element depth, c-DIX performs better when the depth of a given query is low.

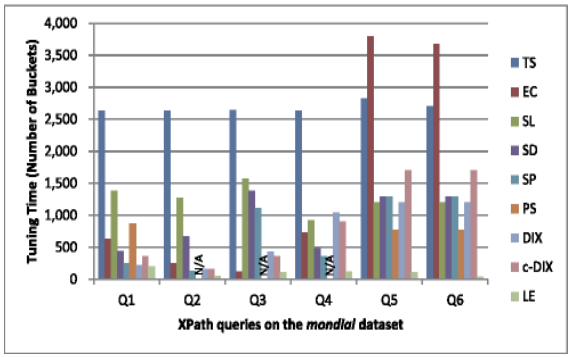

Fig 9: Average access Time Evaluation

V. CONCLUSION

Twig pattern queries enclosing rare statements are popular and critical in XML query processing. In this paper, we introduce a sufficient wireless XML streaming method supporting twig pattern queries. The earlier work on cellular XML streaming only addressed handling of easy path queries. Thus, they are inefficient for twig pattern queries. In contrast, our scheme provides energy and latency is sufficient way to solve predicates and twig pattern matching. Specifically, our scheme decrease the size of the XML stream, exploiting the benefits of the structure indexing and attribute summarization. In addition, our scheme reduces the tuning time as it provides an effective way for selective access of XML elements as well as their attribute values and texts.

VI. REFERENCES

[1] S. Al-Khalifa, H.V. Jagadish, N. Koudas, J.M. Patel, D. Srivastava, and Y. Wu, "Structural Joins: A Primitive for Efficient XML Query Pattern Matching," Proc. Int'l Conf. Data Eng. (ICDE), pp. 141-152, Feb. 2002.

[2] M. Altinel and M. Franklin, "Efficient Filtering of XML Documents for Selective Dissemination of Information," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 53-64, 2000.

[3] S. Amer-Yahia, S. Cho, L.V.S. Lakshmanan, and D. Srivastava, "Minimization of Tree Pattern Queries," Proc. ACM SIGMOD Int'lConf. Management of Data Conf., pp. 497-508, 2001.

[4] N. Bruno, D. Srivastava, and N. Koudas, "Holistic Twig Joins: Optimal XML Pattern Matching," Proc. ACM SIGMOD Int'l Conf. Management of Data Conf., pp. 310-321, 2002.

[5] S. Chen, H. Li, J. Tatemura, W. Hsiung, D. Agrawal, and K.S. Candan, "Scalable Filtering of Multiple Generalizaed-Tree-Pattern Queries over XML Streams," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 12, pp. 1627-1640, Dec. 2008.

[6] C. Chung, J. Min, and K. Shim, "APEX: An Adaptive Path Index for XML Data," Proc. ACM SIGMOD Int'l Conf. Management of Data Conf., June 2002.

[7] Y.D. Chung, S. Yoo, and M.H. Kim, "Energy- and Latency- Efficient Processing of Full- Text Searches on a Wireless Broadcast Stream," IEEE Trans. Knowledge and Data Eng., vol. 22, no. 2, pp. 207-218, Feb. 2010.

[8] D. Florescu and D. Kossmann, "Storing and Querying XML Data Using an RDBMS," IEEE Data Eng. Bull., vol. 22, no. 3, pp. 27-34, Mar. 1999.

[9] T. Imielinski, S. Viswanathan, and B. Badrinath, "Energy Efficient Indexing on Air," Proc. ACM SIGMOD Int'l Conf. Management of Data Conf., pp. 25-36, 1994.

ISSN: 2278 – 1323

[10] H. Jiang, W. Wang, H. Lu, and J. Yu, "Holistic Twig Joins on Indexed XML Documents," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 273-284, 2003.
[11] H. Jiang, H. Lu, and W. Wang, "Efficient Processing of XML Twig Queries with OR-Predicates," Proc. ACM SIGMOD Int'l Management of Data Conf., pp. 59-70, June 2004.
[12] R. Kaushik, P. Bohannon, J.F. Naughton, and H.F. Korth, "Covering Indexes for Branching Path Queries," Proc. ACM SIGMOD Int'l Management of Data Conf., pp. 133-144, June 2002.
[13] C.-S. Park, C.S. Kim, and Y.D. Chung, "Efficient Stream Organization for Wireless Broadcasting of Xml Data," Proc. Int'l Conf. Asian Computing Science Conf., pp. 223-235, 2005.
[14] J.P. Park, C.-S. Park, and Y.D. Chung, "Attribute Summarization: A Technique for Wireless XML Streaming," Proc. Int'l Conf. Interaction Sciences, pp. 492-496, Dec. 2009.
[15] S.H. Park, J.H. Choi, and S. Lee, "An Effective, Efficient XML Data Broadcasting Method in Mobile Wireless Network," Proc. 17th Int'l Conf. Database and Expert Systems Applications (DEXA), pp. 358-367, 2006.
[16] W. Wang, H. Wang, H. Lu, X. Lin, and J. Li, "Effcient Processing of XML Path Queries Using the Disk-Based F&B Index," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 145-156, 2005.
[17] J.P. Park, C.-S. Park, and Y.D. Chung, "Energy and Latency Efficient Access of Wireless XML Stream," J. Database Management, vol. 21, no. 1, pp. 58-79, 2010.
[18] A. Berglund, S. Boag, D. Chamberlin, M.F. Fernandez, M. Kay, J. Robie, and J. Simeon, "XML Path Language (XPath) 2.0," Technical Report W3C, 2002.
[19] Y. Diao, M. Altinel, M. Franklin, H. Zhang, and P.M. Fischer, "Path Sharing and Predicate Evaluation for High-Performance XML Filtering," ACM Trans. Database Systems, vol. 28, no. 4, pp. 467-516, 2003.
[20] M. Francechet, "XPathMark: An XPath Benchmark for XMark Generated Data," Proc. Third Int'l Conf. Database and XML Technologies (XSYM), 2005.
[21] A. Gupta and D. Suciu, "Stream Processing of XPath Queries with Predicates," Proc. ACM SIGMOD Int'l Management of Data Conf., pp. 419-430, 2003.
[22] R. Kaushik, R. Krishnamurthy, J.F. Naughton, and R. Ramakrishnan, "On the Integration of Structure Indexes and Inverted Lists," Proc. ACM SIGMOD Int'l Management of Data Conf., June 2004.
[23] T. Imielinski, S. Viswanathan, and B. Badrinath, "Data on Air: Organization and Access," IEEE Trans. Knowledge and Data Eng., vol. 9, no. 3, pp. 353-372, June 1997.

U.Ravindran received the B.E degree in Computer Science Engineering from Anna University in 2006 and the M.Tech degree in Computer Science Engineering from Bharath University in 2011.Currently, He is an Assistant Professor in the Computer Science Department, Arulmigu Meenakshi Amman College of Engg, Thiruvannamalai Dt, Near Kanchipuram,Anna University, India. ph-9840096489.

SriBhavani.J is currently pursuing M.E degree in Computer Science Engineering, Arulmigu Meenakshi Amman College of Engg, Thiruvannamalai Dt, Near Kanchipuram, Anna University, India.Ph-9952505842.

ISSN: 2278 – 1323