IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014
ISSN: 2320 – 8791 (Impact Factor: 1.479)
www.ijreat.org

XML Automation Tool: A Prime Method for Streaming XML in Wireless Mobile Networks

J.Briskilal[1], Dr. Divya Satish[2]

[1, 2]Department of Computer Science and Engineering, S.K.R Engineering College, Chennai, Tamil Nadu, India

Abstract

Recently, Wireless communications are very prevalent, In order to provide an effective xml streaming for energy and latency efficient by means of Lineage Encoding and Twig pattern queries. Lineage encoding is an active converting scheme which converts byte formats into bit formats thereby providing effective utilization of bandwidth. Also it is used to support twig pattern queries. Twig pattern queries perform multiple way of searching including three tree traversal phases in order to provide a quick reply to the users.xml streaming is done by an innovative method called G node method which includes the combination of multiple elements into a single node. Thereby it provides exact information to the users. We propose an XML automation tool which creates customized xml files .so that there is no need of relying on third party for xml files.. Also Dynamic addition of G nodes is possible in order to add dynamic events without disturbing an existing broadcasting channel.

Keywords: Twig Queries, Dynamic G nodes, XML Automation Tool, Streaming

1. Introduction

With the rapid development of wireless network technologies, wireless mobile computing has become popular. Users communicate in the wireless mobile environment using their mobile devices such as smart phones and laptops while they are moving [1].

Fig.1 shows in wireless XML broadcasting, the broadcast server retrieves XML data to be disseminated from the XML repository. Then, it parses and generates a wireless XML stream. The XML stream is continuously disseminated via a broadcast channel. In the client-side, if a query is issued by the mobile client, the mobile client tunes in to the broadcast channel and selectively downloads the XML stream for query processing. 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.

We need to consider energy conservation of mobile clients when disseminating data in the wireless mobile Environment, because they use mobile devices with limited battery-power. (i.e., energy-efficiency). The overall query processing time must also be minimized to provide fast response to the users (i.e., latency-efficiency). To measure the energy-efficiency and latency-efficiency in wireless broad- casting, the tuning time and access time are used, respectively, [2], [3], [4]. The tuning time is the sum of the elapsed times spent by a mobile client to download the required data. When a mobile client downloads the data (i.e., in the active mode) it consumes more energy than when it waits for data (i.e., in the doze mode). Thus, the tuning time is used as a performance measure for energy-efficiency. The access time is the time elapsed from when a mobile client tunes in to the broadcast channel to when the desired data is completely retrieved from the stream. It is used as a performance measure for latency-efficiency. In Fig. 1, assuming that I is an index segment over the stream and En is the target data, the tuning time is the sum of t1, t2, and t3, whereas the access time is t4.

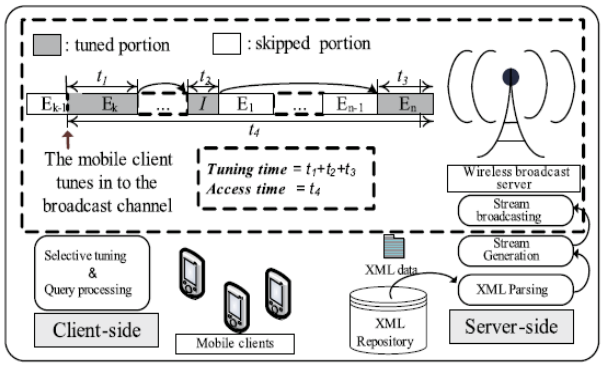

Fig 1.Architecture of Wireless XML broadcasting system

For providing energy-efficient query processing over XML data in wireless and mobile environments, several approaches exploiting the benefits of wireless broadcasting

have been proposed to reduce structural overheads of the original XML document and attach indices containing time information to the XML data stream [5], [6], [7], [8]. We refer to these approaches as wireless XML streaming compared to conventional streaming and processing of XML data in the wired environment. These works enable mobile clients to selectively download the data of their interests by using indices. However, they cannot process XML twig pattern queries efficiently since they do not contain branching information or parent-child relationships

2. Related Works

More than a few researches have been suggested for efficient process of wireless xml stream. S-node [5] creates a unit of xml data stream for wireless xml broadcasting and this process constructs indices by obtaining structural information of the xml document. S-node skips all irrelevant information while the process of downloading. But this process does not support complex twig pattern matching and predicate matching.

XFilter [9] is called xml document filtering system which transforms all XPath queries into a finite state illustration. But this method centers only on filtering time. TwigStack [10] decreases the amount of the midway results and computational cost for integrating the midway results using a chain of linked stacks that represent partial results to root-to-leaf query path.

XR Twig [11] establishes higher performance because it avoids elements which not tally with given twig patterns. G-node [1] a novel unit structure is used to generate an XML stream for Broadcasting. By take advantage of the index information by using Structure Indexing and Attribute Summarization [6], the mobile client can access the stream with shorter latency. This method focus only on static events of broadcasting, also for every XML document, it has to rely on the third party repositories.

3. System Architecture

Fig.2 shows the wireless XML broadcasting; in that the xml file is produced by an automation tool and placed in repository then the broadcast server retrieves XML data from the XML repository for the distribution process. Then that XML file is parsed by the SAX parser and it produces a wireless XML stream.

The XML stream is continuously distribute through a broadcast channel. At the client-side, if a query is delivered by the mobile client, the mobile see through the broadcast station and it downloads the XML stream for query processing.

Also the XML broadcasting structure can support dynamic distribution of G-node without interrupting an existing stream.

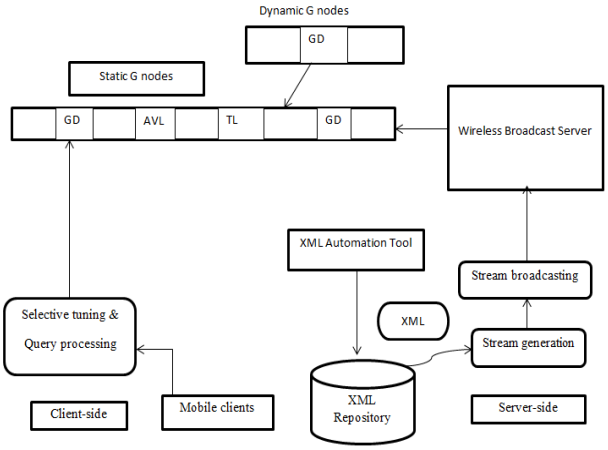

Fig.2 Architecture of Wireless XML Streaming structure

3.1 Working Process of XML Automation Tool

By using XML automation tool custom-built XML file can be generated in a large amount, so that there is no need of depending on third party repositories.XML Automation tool consists of text fields contains root name, add child, add attribute and sub child by using these text fields xml document can be easily generated by submitting the elements by submit button.

Fig.2 shows an XML document which describes a Mondial in turn it consists of countries ,provinces and cities .XML file which is produced from the automation tool is structured, ordered and rooted tree that consists of elements having parent child link.

And this xml file is retrieved from the broadcasting server and SAX parsing will be done. SAX invokes content handlers during the parsing of an xml document.

This automation tool adds an advantage and it avoids the problem of searching the files from third party repositories. After XML file is parsed from SAX parser it is now processed by structure indexing method.

```
<mondial>
    <country id="f0_162" name="Belgium" capital="f0_1477" population="10170241">
        <name>Belgium</name>
        <province id="f0_17457" name="Antwerp">1610695</province>
        <province id="f0_17462" name="Hainaut">
            <located at>West of Belgium</located at>
            <city id="f0_2335" country="f0_162" province="f0_17462">Charleroi</city>
                <population> 255634 </population>
            <city id="f0_2345" country="f0_162" province="f0_17462">Mons</city>
                <population> 188639 </population>
            <city id="f0_2345" country="f0_162" province="f0_17462">Liege</city>
        </province>
    </country>
    <country id="f0_174" name="Bulgaria" capital="f0_1487" population="8612757">BU</country>
    <country id="f0_208" name="Finland" capital="f0_1507" population="5105230">
        <province id="f0_33615" name="Aland">
            <city id="f0_35399" country="f0_208" province="f0_33615">Mariehamn</city>
                <population> 315630 </population>
        </province>
        <province id="f0_33620" name="Haeme">662000</province>
    </country>
    <country id="f0_184" name="Czech Republic" capital="f0_1493" population="10321120">
        <province id="f0_17473" name="Jihomoravsky">
            <located at>East Cost of Czech</located at>
            <city id="f0_2394" country="f0_184" province="f0_17473">Zlin</city>
        </province>
        <province id="f0_17475" name="Severocesky"></province>
    </country>
</mondial>
```

Fig.3 Example xml document

3.2 Lineage Encoding and Attribute Summarization

The Novel algorithm is divided into two main phases:

1. Lineage Encoding, to support queries involving predicates and twig pattern matching. In the proposed scheme, two kinds of lineage codes, i.e., vertical code denoted by Lineage Code (V) and horizontal code denoted by Lineage Code (H), are used to represent parent-child relationships among XML elements in two G-nodes. We propose a light-weight encoding scheme, called Lineage Encoding, to represent parent-child relationships among XML elements in the G-nodes. We also define relevant operators and functions that exploit bit-wise operations on the lineage codes. To the best of our knowledge, our scheme is the first wireless XML streaming approach that completely supports twig pattern query processing in the wireless broadcast environment.
2. 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.

For example, when the Lineage encoding method is applied to the above example document. XML document consists of country, province and city. Lineage encoding works like providing bits of 1 or 0 depending the presence and absence of link between country, province and cities.

Example 1: Country to Province:

Country 1,3and 4 consists of two provinces and country 2 consists of not any province. Hence the Lineage code for country to province is:

LC (V) = 1011
LC (H) = (2, 2, 2)

Lineage code vertical LC(V) represents the link present between country to province and Lineage code horizontal LC(H) represents number of provinces which is present under country.

Example 2: Province to city:

Province 2 consists of three cities and province 3 and 5 consists of one city each. Hence the Lineage code for province to city is:

LC (V) = 011010
LC (H) = (3, 1, 1)

Lineage code vertical LC(V) represents the link present between province to city and Lineage code horizontal LC(H) represents number of cities which is present under province.

Elements in an xml file consists of many attributes of having same name, hence the method called Attribute summarization is used which eliminates the redundant attributes and combine its name and id value pair in single attribute.

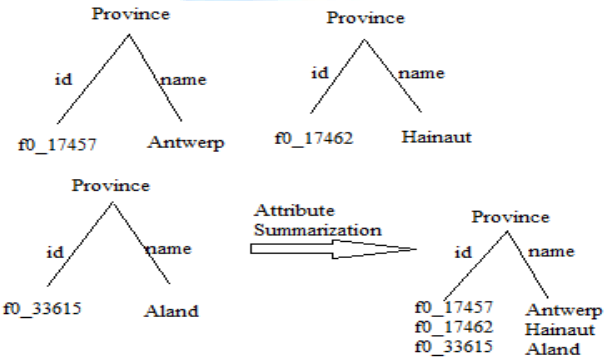

Fig.4 Attribute Summarization of xml document

Fig.3 shows Attribute summarization of the above xml document. Attribute summarization methodology summarizes all the attributes into one single node thereby it reduces the size of the xml stream. Hence the size of the XML stream gets reduced.

3.3 G node and XML Dissemination

We define a streaming unit of a wireless XML stream, called G-node[1]. The G-node structure eliminates structural overheads of XML documents, and enables mobile clients to skip downloading of irrelevant data during query processing. The group descriptor is a collection of indices for selective access of a wireless XML stream. Node name is the tag name of integrated elements, and Location path is an XPath expression of integrated elements from the root node to the element node in the document tree.

All the group nodes are broadcasted through Wi-Fi devices that are received by the mobile devices'

3.4 Query Tree Formation and Selective Tuning

In this section, we describe how a mobile client can retrieve the data of its interests. Assuming that there is no descendant axis in the user query, query processing algorithms for a simple path query and a twig pattern query are presented. when the current node is the leaf node it chooses normal query processing otherwise it chooses twig pattern query matching [12].

Example:

Q1: //Country [name/text () ="Belgium"]/province/city

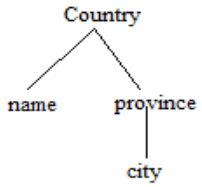

Fig.5 Sample query tree with twig pattern query

Q2: //Country/Province[located at/text () ="Middle"]/city

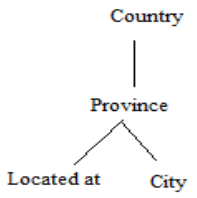

Fig.6 Sample query tree with complex twig pattern query

Fig.5 and Fig.6 represents sample query tree with normal and complex twig pattern query representation. In simple normal query processing tree traversal is done in depth first traversal process. In twig pattern query processing since the predicate condition is too complex the query tree which is formed is complex branched, hence the tree traversal is done in three phases of traversal.

The three tree traversal phases are Tree traversal phase, Sub path traversal and Main path traversal. In Tree traversal phase traversal is performed in depth first approach. In Sub path tree traversal phase traversal is performed at the highest branch of the query tree and it performs the depth first traversal. In the main path traversal it traverses from the root to leaf node of the query tree. This twig pattern query processing provides very quick reply to the users, so that latency is reduced.

3.5 Dynamic G nodes

Dynamic addition of GNode ensures the credibility of the Broadcast system efficiently proposed by our approach. AVL tree and Structured Indexing process will be handled that will probably affect the XML document in temporary buffer. Dynamic modification of Attribute value enables to change any data on the broadcast stream whenever needed and is achieved by the Attribute summarization mechanisms and the Structured Indexing of XML data handled in our system.

4. Implementation

Implementation is performed in Java platform jdk environment at the server side. And at the client side mobile device which supports of Android OS of any version and for broadcasting Wi-Fi router is included.

Initially an xml document is created using XML Automation tool and it is placed in an xml repository thereby relying on third parties is reduced. At the server side when the above Xml document is extracted from the repository it performs structure indexing and Lineage encoding.

When the above xml document is processed by structure indexing and Lineage encoding, it captures the structure information and provides indices according to that and provides bit formats according to the links.

Here the example is performed in people XML document which is generated by XML Automation Tool.

People XML document consists of name, type and age of the people.

```
<People>
<Type> male
</Type>
<Name> aaa
<Name>
<Age> 5
</Age>
<Type> female
</Type>
<Name> bbb
</Name>
<Age> 5
</Age>
<People>
```

Fig Sample people xml document

Structure Indexing:

{People=1, Type=4, Name=7, Age=4}

Lineage Encoding:

{People-Type=1-4, People-Age=1-4, Age-Name=0100-1, Age-Age=0100-1, Age-Type=0100-1, Type-Name=11-00, Name-Age=011-00, People-Name=1-7}

G nodes:

Type -> [NodeName=Type, LocationPath=People/Type, CI={Age=5#5, Name=aaa#bbb, Type=male#female}, LC(V,H)=1-4, AVL=null, TL=null]name -> [NodeName=Name, LocationPath=/People/Age/Name, CI={Age=5#5, Name=aaa#bbb, Type=male#female}, LC(V,H)=11-00, AVL=null, TL=null]Age -> [NodeName=Age, LocationPath=/People/Age/Name/Age, CI=Age=5#5, Name=aaa#bbb, Type=male#female}, LC(V,H)=011-00, AVL=null, TL=null]

To extract the data when the query is issued it retrieves the exact information of the given sample people document.

5. Conclusion

In this paper, an efficient wireless XML streaming method supporting Twig pattern queries is proposed. This scheme provides an energy and latency efficient way to evaluate predicates and twig pattern matching. Specifically it reduces the size of the XML stream, exploiting the benefits of the structure indexing and attributes summarization.

In the future, we plan to analyze the creation of XML using an XML Automation tool to generate the xml and send the data directly for processing without keeping in the repository in a depth full manner. Also we plan to analyze the timing of dynamically choosing pattern so that still the latency gets reduced. To, analyze dynamically updating the data in the xml document.

Acknowledgements

At the very outset, we wish to express our thanks to all those who were involved in the completion of this work. We express our immense thanks to Dr. Suguna, Professor and Dean of Department of Computer Science and Engineering, SKR Engineering College for giving us support and encouragement for carrying out this project. We also express our thanks to all of our family and friends.
.

References

[1] J.P. Park, C.-S. Park, and Y.D. Chung "Lineage Encoding: An Efficient XML Streaming Supporting Twig Pattern Queries," IEEE Trans. Knowledge and Data Eng., vol. 25, no. 7, July 2013.

[2] 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. Gregory Blanc,Ruo Ando and Youki Kadobayashi "Term Rewriting Deobfuscation for static Client Side Scripting Malware Detection"IEEE 2010.

[3] 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. Sanctum Inc. "Wen Application Security Testing – Appscan 3.5", http://www.sanctuminc.com.

[4] 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. J.Adam Kiezun and Karthick Jayaraman "Automatic Creation of Sql Injection and XSS Attacks" ICSE'09 May 16-24,2009.

[5] 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. Fang yu,Muath alkhalaf and Tevfi Bultan "Stranger :an automata-based string analysis tool for php" this work is supported by NSF grants.

[6] 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.Gary wassermann and Zhendong su "Static Detection of Cross site scripting Vulnerability"University of California,IEEE 2010.

[7] 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. Siddharth Tiwari ,Richa Bansal "Optimised client side solution for Cross site scripting" IEEE 2008.

[8] 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. Mozilla Foundation, http://www.mozilla.org,2008.

[9] M. Altinel and M. Franklin, "Efficient Filtering of XML Docu- ments for Selective Dissemination of Information," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 53-64, 2000. Cygwin.Cygwin Information and Installation. http://cygwin.com/,2008.

[10] 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.

[11] 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.

[12] 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.