Chapter 18
Experiments and their Assessment

Ibrahim Dweib
Sultan Qaboos University, Oman

Joan Lu
University of Huddersfield, UK

ABSTRACT

In this chapter, the authors give a description of the experiment setup consisting of experiment environment and performance measurement. They perform experiments on mapping XML document into relational database, building XML document from relational database, updating XML document stored in relational database, and retrieving document content from relational database using XPath expressions. These experiments will be done to check the scalability and effectiveness of the model. Then they compare their model with the (Tatarinov et al., 2002) and the Accelerating XPath model (Torsten et al., 2004). The comparison consist of four stages: mapping, building, updating, and retrieving, as most of other studies just took one or two stage and forgot the others. Some of them took retrieving, others took updating, and others took updating and retrieving, but most of them did not consider mapping and rebuilding.

EXPERIMENT SETUP

Experiment Environment

All experiments tests are conducted on a PC of an Intel Core2 Quad Q9550 2.83 GHz CPU, 4.00 GB RAM, running Windows 7 Professional. Visual Basic 6 programming language is used to implement MAXDOR model, and Microsoft Access 2007 is used as a target relational database for storing XML document contents on local hard drive. In addition, a disk file is named with document number in the document table and with an XML extension created for reconstructed XML document from relational database.

DOI: 10.4018/978-1-4666-1975-3.ch018
Performance Measurement

- Mapping XML document into RDB execution time.
- Rebuilding of XML document from RDB execution time.
- Dealing with any document size.
- Inserting nodes processing time (number of nodes to be relabelled).
- Query processing execution time.

The execution time is used as an evaluation scale in this research rather than storage space since the former is crucial nowadays for the users, while storage space is available in a very huge size with reasonable prices.

TESTING STRATEGIES

Mapping XML Document into Relational Database Performance

The experiment is performed as follows:

Face 1, scalability test: An XML document generator from XMark (Busse et al., 2002) is used to create documents of different sizes with factors of 0.1, 0.2, 0.3, 0.4 and 0.5. The documents characteristics are shown in Table 1. In this experiment, our model shows performance in a linear and scalable manner as document size is increasing. The mapping result over different sizes of the same document is shown in Figure 1.

   Face 2, effectiveness test: Three groups of documents of different sizes 11MB, 82MB and 107MB but with different structure and different numbers of token are included in this experiment. Table 2 shows documents properties and their mapping and rebuilding time. Figure 2 shows the time required for mapping XML documents into relational database which consistently increases as the number of tokens increases in the document.

   Considering the results shown in Figure 1 for homogenous documents and those shown in Table 2 and Figure 2 for heterogeneous documents coupled with calculating the correlation coefficient between document size and mapping time in the two cases $r_1=0.99988$ and $r_2=0.8751$ on the one hand, and the number of tokens and mapping time in the two cases $r_3=0.99991$ and $r_4=0.9991$ on the other hand, we can conclude that the time required for mapping the document largely depends on the number of tokens (i.e., elements and attributes) in the document, the document size and document depth ($r=0.1752$) respectively.

Now let us compare MAXDOR model with Global Encoding for (Tatarinov et al., 2002) and Accelerating XPath for (Torsten et al., 2004), since the three models are using the same general number encoding to identify the XML document of elements and attributes (tokens). A detail description for Global Encoding and Accelerating XPath is given in Chapter 15.

   The three models use one scan to shred the document contents, assign an identifier for each token, reserve node information, (i.e. token name and token value) to store them in one tuple in relational database. Global Encoding adds another table for tokens path from the document element

<table>
<thead>
<tr>
<th>Document Name</th>
<th>Factor used</th>
<th>Document Size (MB)</th>
<th># of nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auction_1</td>
<td>0.1</td>
<td>11.3</td>
<td>206130</td>
</tr>
<tr>
<td>Auction_2</td>
<td>0.2</td>
<td>22.8</td>
<td>413111</td>
</tr>
<tr>
<td>Auction_3</td>
<td>0.3</td>
<td>34.0</td>
<td>616229</td>
</tr>
<tr>
<td>Auction_4</td>
<td>0.4</td>
<td>45.3</td>
<td>820438</td>
</tr>
<tr>
<td>Auction_5</td>
<td>0.5</td>
<td>56.2</td>
<td>1024073</td>
</tr>
</tbody>
</table>
Related Content

Data Storages in Wireless Sensor Networks to Deal With Disaster Management

Towards a Statistical Approach to the Analysis, the Indexing, and the Semantic Search of Medical Videoconferences

Schema Independent XML Compressor
[www.igi-global.com/chapter/schema-independent-xml-compressor/75903?camid=4v1a](www.igi-global.com/chapter/schema-independent-xml-compressor/75903?camid=4v1a)

Probabilistic Approach to Arabic Speech Correction for Peoples with Language Disabilities
[www.igi-global.com/article/probabilistic-approach-to-arabic-speech-correction-for-peoples-with-language-disabilities/132488?camid=4v1a](www.igi-global.com/article/probabilistic-approach-to-arabic-speech-correction-for-peoples-with-language-disabilities/132488?camid=4v1a)