

User's Guide

Benchmarking Summarizability Processing in XML Warehouses with Complex Hierarchies

By Chantola KIT, Marouane Hachicha, and Jérôme Darmont

first.last@univ-lyon2.fr

The program is run with Java NetBeans (JDK 1.7) to generate XML data with complex hierarchies, QBS, and Pedersen queries.

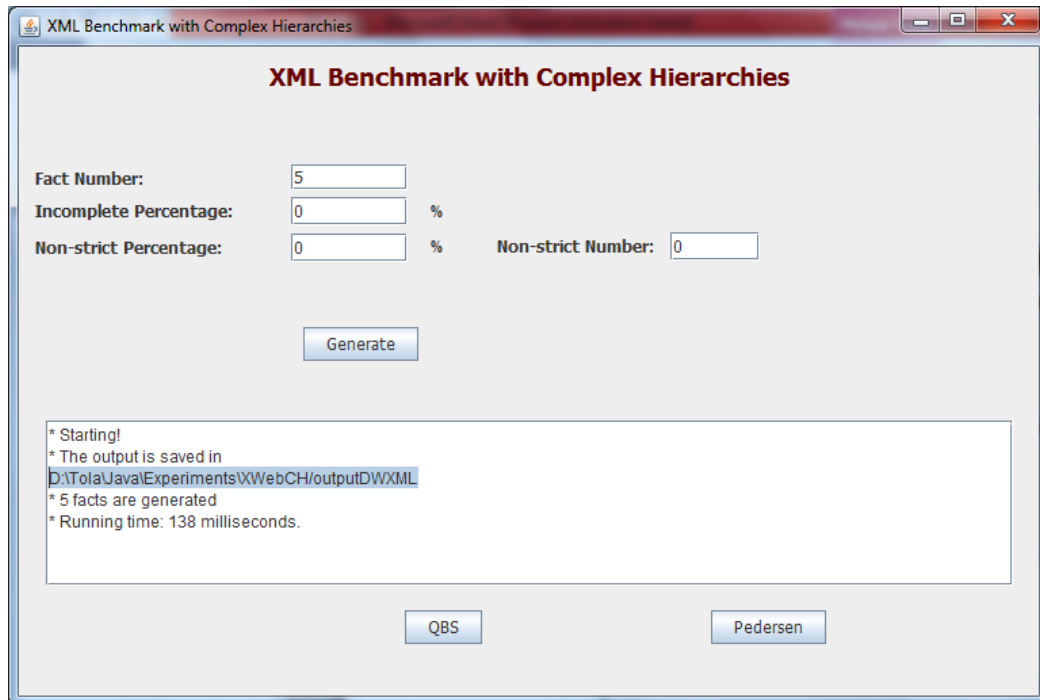
I. Source Code

The source files are organized into directories:

- src/ contains the Java code of the application organized in packages;
- lib/ contains the *.jar files needed for program execution;
- data/ contains XWeB (TPC-H/dbgen) source data files;
- dimensionXML/ contains XWeB dimensions in XML format
- outputDWXML/ is a temporary directory for storing generated fact and dimension XML documents
- PedersenOverheadXML/ stores summarizable XML documents produced before query processing of Pedersen
- outputGrouping/ store XML documents which are the result of QBS and Pedersen grouping

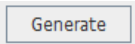
II. Generate complex hierarchy XML data warehouse

1. Open java project XWebCH
2. In XWCH package, double click Run.bat (to run `.\XWebCH\src\XWCH\MainForm.java`)
3. You will see the form as follow

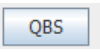


The screenshot shows a Java Swing window titled "XML Benchmark with Complex Hierarchies". The window has a title bar with standard Windows controls (minimize, maximize, close). The main content area has a light gray background. At the top, the title "XML Benchmark with Complex Hierarchies" is displayed in a bold, dark red font. Below the title, there are four input fields with labels: "Fact Number:" with a text box containing "5", "Incomplete Percentage:" with a text box containing "0" and a "%" symbol to its right, "Non-strict Percentage:" with a text box containing "0" and a "%" symbol to its right, and "Non-strict Number:" with a text box containing "0". Below these fields is a "Generate" button. At the bottom of the window, there are two more buttons: "QBS" and "Pedersen". A text area at the bottom of the window contains the following text: "* Starting!", "* The output is saved in", "D:\TolaJava\Experiments\XWebCH\outputDWXML", "* 5 facts are generated", and "* Running time: 138 milliseconds."

Figure 1. MainForm

4. Set the values of fact number, incomplete percent, non-strict percent, and non-strict number
Note:
fact number > 0
incomplete percentage ≥ 0
non-strict percent ≥ 0
non-strict number > 1 if non-strict percent > 0
5. Click on  button to generate complex hierarchy XML Data warehouse which is stored in the specified directory (`.\XWebCH\outputDWXML`)

III. QBS Query

1. On Figure 1, click on  button, you will see the window below.

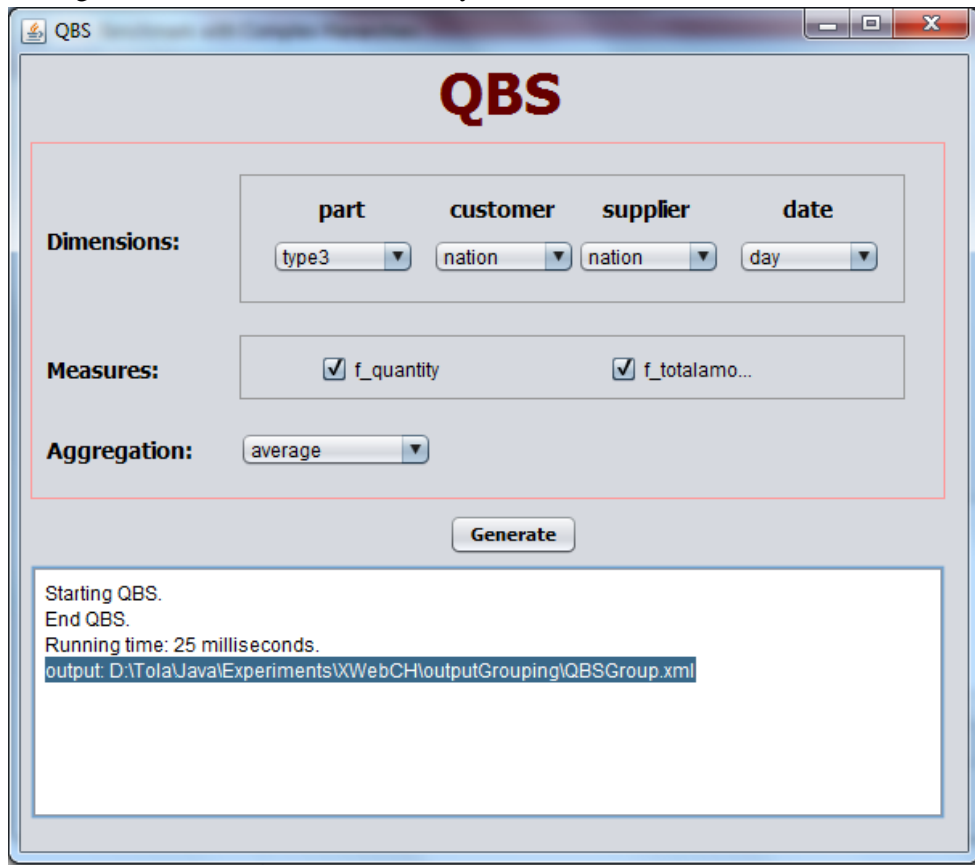
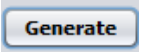


Figure 2. QBS Query

2. In Figure 2, you can make QBS query by specifying:
 - a. Dimensions for multidimensional cube
 - b. Measures
 - c. Aggregation

Note: at least one dimension and one measure are selected

The current version, for ease of use, we use Java GUI for specifying the query. In the future we will include XQuery

3. Click on  button to run the query
4. The output will be an XML document whose file location is specified in the output text (.\XWebCH\outputGrouping\QBSGrouping.xml).

IV. Pedersen Query

1. On Figure 1, click on  button, you will see the window below.

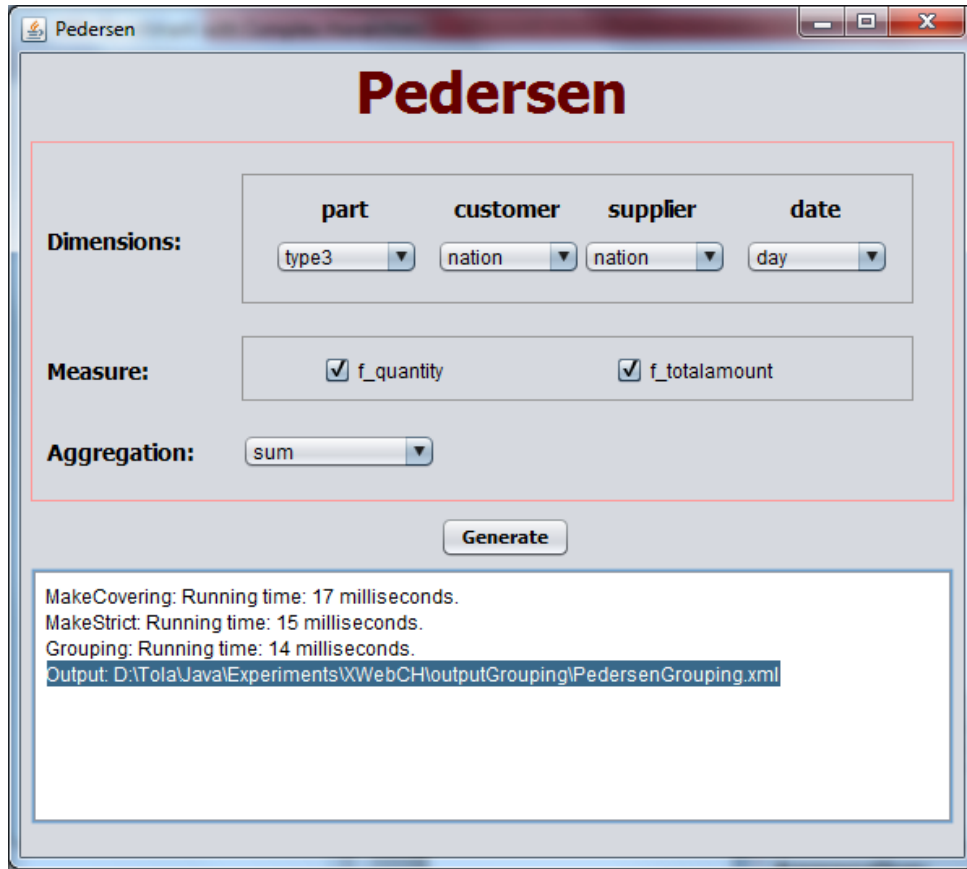
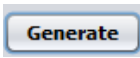


Figure 3. Pedersen Query

2. In Figure 3, you can make Pedersen query by specifying:
 - a. Dimensions for multidimensional cube
 - b. Measures
 - c. Aggregation

Note: at least one dimension and one measure are selected

The current version, for ease of use, we use Java GUI for specifying the query. In the future we will include XQuery

3. Click on  button to run the query
4. The output will be an XML document whose file location is specified in the output text (.\XWebCH\outputGrouping\PedersenGrouping.xml).

-----The end-----