CSE - 4255 Data Mining and Warehousing Lab
Data Warehousing

Saif Mahmud   M. Tanjid Hasan Tonmoy
Roll: SH - 54   Roll: SH - 09

Submitted To:
Dr. Chowdhury Farhan Ahmed
Professor

&

Abu Ahmed Ferdaus
Associate Professor

Department of Computer Science and Engineering
University of Dhaka

November 11, 2019

1   Problem Definition

The tasks for this assignment is described below:

- Creating a Relational Database

- Defining Schema (Star, Snowflake or Galaxy) for Data Warehouse

- Creating Dimension Table and Fact Table from Predefined Relational Database

- Creating Data Cuboid from Defined Schema

- OLAP operation on the data cube such as roll-up, drill-down etc.

2   Theory

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of managements decision making process. A data cube allows data to be modeled and viewed in multiple dimensions. It is defined by dimensions and facts.

The most common modeling paradigm is the star schema, in which the data warehouse contains (1) a large central table (fact table) containing the bulk of the data, with no redundancy, and (2) a set of smaller attendant tables (dimension tables), one for each dimension. The schema graph resembles a starburst, with the dimension tables displayed in a radial pattern around the central fact table.

The snowflake schema is a variant of the star schema model, where some dimension tables are normalized, thereby further splitting the data into additional tables. The resulting schema graph forms a shape similar to a snowflake.

Sophisticated applications may require multiple fact tables to share dimension tables. This kind of schema can be viewed as a collection of stars, and hence is called a galaxy schema or a fact constellation.

The roll-up operation (also called the drill-up operation by some vendors) performs aggregation on a data cube, either by climbing up a concept hierarchy for a dimension or by dimension reduction.

Drill-down is the reverse of roll-up. It navigates from less detailed data to more detailed data. Drill-down can be realized by either stepping down a concept hierarchy for a dimension or introducing additional dimensions.

3 Conclusion