

IMPROVED DATA MODEL WITH REFINED STRUCTURES

Improved Data Model With Refined Structures For Course Selection And Opinion Satisfaction In E- Learning

Atika Qazi

atikaqazi@yahoo.com

Ram Gopal Raj,

Mehwish Waheed⁴

University of Malaya, Kuala Lumpur

Syed Ghayur Naqvi

Atika Qazi

COMSATS University, Pakistan

This paper has been presented in the



organized by

School of Business and Economics

University of Management and Technology, Lahore, Pakistan

This paper has been included in the conference proceedings with good intentions, where the conference and its organizers are not liable at all for the contents of this paper and / or any part of it. For more information about the conference please visit the conference website: <http://cgr.umt.edu.pk/icobm2013/index.html> or write the organizers at icobm@umt.edu.pk

ABSTRACT

The success of every enterprise depends upon the decisions made by decision makers. Data warehouse is novel concept and plays an important role for decision making by the use of OLAP. There are diverse fields to select and study by the students in this modern era. To follow such a dimensional analysis approach that is more beneficial for selecting the curriculum by students and the multi-dimensional cube analysis takes place with more refined and elaborated ways. The data mining and OLAP techniques are also applied to prepare the standard platform to analyze the query by multidimensional view. To analyze the curriculum by multiple angles ultimately serves the university teaching decision support system. Before settled down dimensions a survey may be conducted to analyze the trends and demand, in order to make a successful data model. The case study at university level is taken to design the views more advantageous manner. The data is gathered from multiple reliable resources and the query is generated in more improved way.

Key words: Online Analytical processing (OLAP), Distributed Data Warehouse (DDW), Data Modeling (DM), Dimensional Modeling (DM)

1. INTRODUCTION

A data warehouse is a logical collection of information gathered from many different operational databases used to make business intelligence that supports business analysis activities and decision making tasks. It is used for providing the basic infrastructure for decision making by Extracting, cleansing and storing huge amount of data. Data warehouses support business decisions by collecting, consolidating, and organizing data for reporting and analysis with tools such as online analytical processing (OLAP) and data mining.

Data warehouse is centralized data repository maintained separately from organization's operational databases to help organization in corporate decision making process. William Inmon has described data warehouse as *"A subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management decisions"* [2] *"Data warehouse is a set of materialized views over data sources"* [4], [5], [6] Ralph Kimball et. Al. defined *"A data warehouse is a copy of transaction data specially structured for query and analysis"* [7].

The requirement of today is how to extract useful information from the bulk of data and what approach might be followed in order to make more reliable decision making to attain maximum advantage of technologies for the success of origination. As the right decision at times leads towards more success and enterprises get beneficial by the decision made by decision makers.

OLAP is a useful tool that supports the complex query analysis and is practice for effect for decision making that is carried on by decision makers. The query is directed by the analysis to OALP and is process query accordingly and once query is processed the result are generated back to the decision maker's men and managers. The generated results against queries are viewed by OLAP that runs on tables.

The following are the research issues

- a. The query must be analyzed thoughtfully by query experts in order to avoid false analysis and save time to increase efficiency.
- b. Query must be more optimized and calculated.
- c. The tables are more organized and well explained in order to gain maximum advantage.
- d. Any new concern is also added as enhancement part.
- e. The tables are designed carefully that it covers maximum of information as the query must be run on tables so tables are better to prepared via little survey. or can say that data that is gathered for blind table must be explanatory.

In this proposed work the more refined and strategic approach is used for the analysis of curriculum chosen by students in a data warehouse environment. The improved structure for data modeling is proposed to increase success ratio and make overall performance efficient and fast. The tables are created after the survey study in a university in order to get the required data that might frequently asked for the development of successful educational environment.

This research work applies the data warehouse technology in analysis data model of curriculum chosen by students from diverse angles from the teachers to choose the course and for the head to allocate the course by taking OLAP, to provide the accurate information to decision makers for the establishment of curriculum and teachers and head allocated tasks by applying query in more generic and optimized manner. The data that is necessary for teachers and students and all parameters that are involved for syllabus and course selection.

2. DESIGN OF DATA WAREHOUSE FOR SELECTION OF COURSE BY RELATED INDIVIDUALS

The data warehouse together with tools such as OALP or data mining are collectively referred as business intelligence technologies. Basic architecture of data warehouse is discussed in Error! Reference source not found.. Connolly et. al. proposed three tier architecture of a data warehouse 3. Atika Qazi et.al proposed the improved architecture for distributed data warehouse environment [1]. Detailed logical architecture of a data warehouse is presented in Error! Reference source not found..

The design of data model for curriculum chosen by students is the most critical step. The dimensionality modeling is used for selection of model. DM is the name of logical design techniques often used for data warehouse. There are usually star schema, snowflake and star flake schema. The case studies follow the star model the logical model that contains the factual data in the center and dimensions tables around it. The fact table contains the numerical measure "facts" and dimension tables, normally contains the textual information. Figure 1 may support a query that requires selecting the curriculum by students and teachers and head for allocation of related work. *"Star schema can be sued to speed up the query performance by deformalizing the related information"* Thomas connolly, data base management system, Fourth Edition.

The dimension is the angle form where user observes data. The dimensions have different according to the need of requirement. Structure of dimensional model is like as.

The first step in the development of structural dimensions is to correspond very closely to what one normally does in relational database. The star model that in the case study is proposed depends upon taking the central intersection entities as the fat tables and building the foreign key equivalent primary key relations as dimensions.

The dimensional modeling is probable. The report writing, query tools and user interfaces can all make strong expectations to make the user interfaces more logical and to make processing more effective.

There are certain steps in the design process

2.1 Choose the data mart

A set of related fact and dimension tables

Single source or multiple sources

Conformed dimension

Typically have a fact table for each process

2.2 Declare the grain

The grain is the unit of analysis that determines what each fact and record represents. For example Individual transactions Snap shots Line items on documents.

2.3 Choose the dimensions

Dimension attributes are used as a source of most of the interesting constraints in data warehouse queries, and they are virtually always the source of the row headers in the SQL answer sets.

2.4 Choose the facts

The primary keys of all the related dimensions are facts and mostly are numeric that uniquely identifies each dimension. Useful facts tend to be numeric and additive.

2.5 Fact and Dimension

The factors are elaborated discussed then dimensions may takes place in proper advantageous from.

2.6 The Survey Study For Creation Of Data Warehouse's Star Model

The survey study covers the facts related dimension tables that leads to successful results are given below:

2.7 What are the factors considered for selecting a course by a student?

2.7.1 Course of interest

It involves what is the choice and interest of student in presented course

2.7.2 Teacher of interest

Teacher may have different strategy of teaching a course, students might have better understanding with one than others.

Trend of course/Market demand/ Job trends

Time to time demand varies in the market according to the requirement of that time. If application of specific course is high in the market the selection of course must be effected with trend.

Fact Table	Dimensional Table	Connections
Key- Uniquely Identifies Record	Keys	Between Dimensions & Facts
Attributes	Attributes

Table 1

2.7.3 Time duration (maximum time to complete the course)

Time factor is involved also in selection of course, few courses takes long time span to complete then others that takes less time to complete.

2.7.4 Social and political climate

The long terms disorder in overall environment may effects education system. The people do go for earning money in any way rather to get admission.

2.7.5 Far reaching approach /Distance

As with time to time the cost of living is becoming higher than yesterday, some people don't afford to pay high rent for coming towards institutes to study rather prefer nearest station to continue with studies.

2.7.6 Course category /lecture mode (traditional /distance learning)

The beginners prefer to follow traditional system and some prefer to distance learning. The situation vary time to time depend upon the convenience of the related person.

2.7.7 Course type (elective /compulsory)

The course may have different types and student have to check for the category of course as well.

2.7.8 Hostel accommodation (allowed not allowed)

The people who belong to far reaching areas must look for hostel accommodation.

2.7.9 Locality approach

Some places are at very busy point and it might create difficulty for the one to come and start work.

2.8 What are the factors to be considered by a teacher for offering a course?

2.8.1 Teacher area of specialization

Tutor prefers to take the course if it's under his/her range.

2.8.2 Teachers work load

The work load of teacher is also important factor in order to manage the allocated class with full attention

2.8.3 Class of interest

Sometimes teachers have taught the same class then if he/she is willing for the same class in coming semester or not.

2.8.4 Strength of class

The total number of students also matter to lead the class at its best.

2.8.5 Maximum duration of course

The time period of related course is also imported factor to take the course.

2.8.6 Availability of basic necessities (equipment's)

The few courses may also involve the lab work and for teaching at its best the practical demonstration is required.

2.8.7 Lectures Timings

Some of the lectures may takes less time to teach some required more time.

2.8.8 Availability (leave period)

The allocated teacher may not be available in the whole semester if he/she applied for the long leave.

2.8.9 Distance at some extent

The distance is also seeing feature to reach the place in time.

2.8.10 Pay packages

The organization may vary with their pay packages and it's also an imported factor to join at offered package or no.

2.8.11 Immediate boss

By selection of the relevant course the related boss is also considered.

2.9 What are the factors considered by HOD for allowing to be offered course?

The head also consider many factors as discussed:

2.9.1 Teacher's availability

The teachers might be in process of long leave and having chances to get approved.

2.9.2 Teachers demand on related course

The area of expert is also preferred in order to allocate the course.

2.9.3 Students demand on related course

The feedback of teacher to allow for more courses is also involved.

2.9.4 Student's strength on that course

Sometimes it's difficult for teachers to manage the big class as compare with small group.

2.9.5 Course prerequisites

The prerequisites are checked before offering related course to students.

2.9.6 Course importance

The compulsory courses are allocated to more experienced teachers.

2.9.7 Availability of course

The availability of course depends on the availability of resources as well.

2.9.8 Teacher work load

The teacher is assigned leveled work in order to balance workload.

Attributes associated to factors are given below in the table:

2.10 What are the key attributes of course?

Course type (Elective /Compulsory), Course prerequisites, Credit hours, Course name, Course code, Department, Cost type (under grad /post grad), Course tutor, Course registration fee, Course offered semester, Total no of lectures for offered course, Required Resources, Lab course, Course related Faculty, Semester , Course offered season.

2.11 Conceptual Model For Dimension Table

The conceptual model includes multiple parameters as discussed below:

2.11.1 What are the key attributes of students?

Student name, Student Id No., Grade, Department, Specialty, Gender, Student DOJ, Student Date of Birth, Student status (under grad /post grad), Student attendance record, Student course duration, Student is on regular fee / others (fee mode regular /other), Students dues paid /pending, Special child /normal , Regular student / Quota system, National student / International student, Student Pass out date.

2.11.2 What are the key attributes of teacher?

Teacher name, Teacher Id, Teacher relevant department, Professional title, Teacher qualification Teacher experience, Teacher category (regular /visiting), Gender, Teacher course, Teacher normal /special, Teacher date of joining, Teacher period of leave, Teacher research publications, Teacher allotted projects, Projects proposed by teacher.

2.11.3 What are the key attributes of semester

Semester ID No, Semester year , Semester season (fall /spring), Semester break session, Semester break duration, Semester of seasonal courses, Semester dues (revised / static)

3. The Dimensional Modeling with Testing and Concealed Approach

The huge data is gathered from different resources and then they are extract transform and load into data warehouse. The filtered data that is stored into data ware house is used for

multi-dimensional analysis of related work, for educational system the education related data for course selection teachers allocation to a course and course assignment to the teachers the data warehouse is upheld and used frequently. The related stored information is used to support decision making and eventually helps in the success of an organization.

The multi-dimensional cube analysis takes place for crucial decisions related to educational institute during course selection period by students and teachers. The OALP is used for analysis and passing queries by teachers and student and any concern and allowed person. The proposed MODEL may uses the Microsoft SQL server to release the star model to transform cube analysis and the survey conducted by experts to settle down the expected attributes and dimensions, also testing may takes place by experts in order to check for the presence of related data if it's sufficient to answer the multiple queries find the errors that might expected and to fix them. The survey process may be little lengthy but once if it is settled then it may help for decision making in long run for other then expected domain, for example the described attributes are not for only the selection of curriculum by students by it also helps in different analysis. The more elaborated attributes and their dimensions the flow chart and assessment table plays an important role in the successful decision making. The proposed model may increase success ratio of related organization.

4. Dimensional Modeling Flow chart

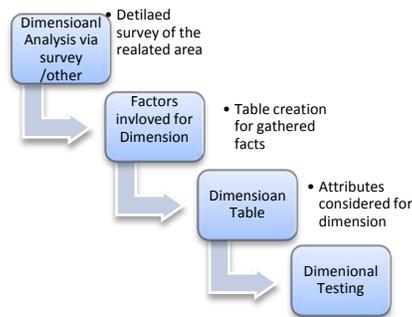


Fig 1 Dimensional Modeling

In the first step the survey study takes place from relevant sources and then tables are created on the bases of collected facts invalid data may discarded then dimension are arranged from facts drawn, finally

the testing takes place to check either the structured model is working at its premium.

Proposed Dimensional Model

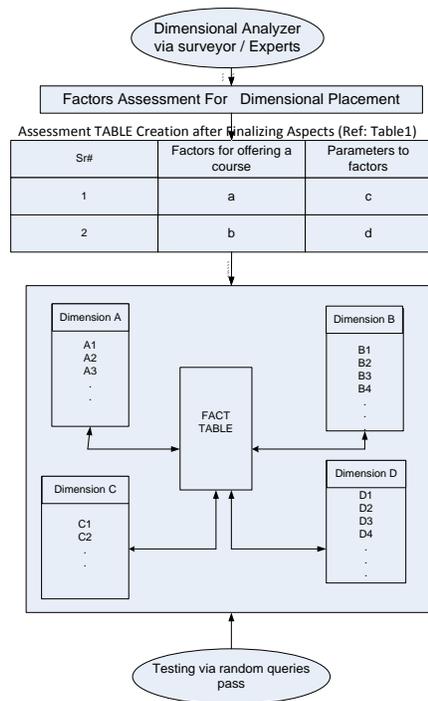


Figure 2 Proposed Dimensional

Models

5 CONCLUSIONS

The proposed model is suitable for the data warehouse analytical demands and it is most significant to work over the related field and then settle down the dimensions and attributes in order to make more reliable decisions. The proposed model is useful not only for curriculum selection but it will fulfill the need of diverse queries related to education setup. The testing after building the model also plays an important role for the successful decision making and leads the success of an organization. The research questions explained above are worked out in future work.

Assessment Table

The Assessment table contains the factors and the parameters against it. Once this table is created then the expert may review it time to time and for future study the same table might be used with more edition if required.

Factors for offering a course	Parameters to factors
1. Factors considering by a Student to select a course.	<ul style="list-style-type: none"> • Course of interest • Teacher of interest • Trend of course/Market demand/ Job trends • Time duration (maximum time to complete the course) • Social and political climate • Far reaching approach /Distance • Course category /lecture mode (traditional /distance learning) • Course type (elective /compulsory) • Hostel accommodation (allowed not allowed) • Locality approach • Course prerequisites • Course importance • Availability of course • Teacher work load
2. Factors Considering By Teacher To Offer A Course	<ul style="list-style-type: none"> • Teacher area of specialization • Teachers work load • Class of interest (if once that class is being taught by same teacher what was experience good /avg/ bad) • Strength of class

	<ul style="list-style-type: none"> • Maximum duration of course • Availability of basic necessities (equipment's • Lectures timings • Availability (leave period) • Distance at some extent • Pay packages • Work load • Immediate boss
<p>3. Factors Considering By HOD To Offer Allocate Course To Teacher</p>	<ul style="list-style-type: none"> • Teacher's availability Teachers demand on related course • Students demand on reacted course • Student's strength on that course

REFERENCES

- [1]. Atika Qazi “A Distributed Data Warehouse Architecture with Fair Query Execution Scheme”,2010 ICMLC Singapore.
- [2]. Thomas Connolly, Carolyn Begg, Database Systems: A Practical Approach to Design, Implementation and Management, 4th Edition, Addison-Wesley, 2003
- [3]. William Inmon, *Building the Data Warehouse*, 2nd Edition, New York: Wiley publisher. Inc, 1996
- [4]. Z. Bellahsene, Schema, “Evolution in Data Warehouses”, *Knowledge and Information Systems*, Springer-Verlag, pp 283-304, 2002
- [5]. E.A. Rundensteiner, A. Koeller, X. Zhang, “Maintaining Data Warehouses over Changing Information Sources”, *Communications of the ACM*, Volume, 43, New York, NY, USA, pp 57-62, 2000
- [6]. Ralph Kimball, M. Joy and T. Warren, *the Data warehouse Toolkit: with SQL server and Microsoft Business Intelligence Toolset*, 2nd Edition, New York: Wiley publisher. Inc, 2006.
- [7]. Efraim Turban, Jay E. Aronson and Narasimha Bolloju, *Decision Support Systems and Intelligent Systems*, 7th edition, Prentice Hall College Div, 2001
- [8]. Mohammad Rifaie, Erwin J. Blas, Abdel Rahman M. Muhsen, Terrance T. H. Mok, Keivan Kianmehr, Reda Alhaji, Mick J. Ridley, “Data warehouse Architecture for GIS Applications”, In Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services (iiWAS '08) , November 2008, Linz, Austria.
- [9]. Li Lin-hua, “The Design of multidimensional analytical model of data warehouse”, *Computer Project and Application*. 20 (2004),pp.30-35.
- [10]. Thanh Binh Nguyen, A. Min Tjoa, Roland Wagner, “Conceptual multidimensional data model based on MetaCube”, *Lecture Notes in Computer Science*. (2000).