Water Science and Engineering, **Sep. 2008, Vol. 1, No. 3, 112-120**
ISSN 1674-2370, ***http://kkb.hhu.edu.cn*, e-mail: *wse@hhu.edu.cn***

Multi-dimensional database design and implementation of dam safety monitoring system

Zhao Erfeng*[1, 2], Wang Yachao[1, 2], Jiang Yufeng[1, 2], Zhang Lei[1, 2], Yu Hong[1, 2]

1. National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing 210098, P. R. China

2. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, P. R. China

Abstract: To improve the effectiveness of dam safety monitoring database systems, the development process of a multi-dimensional conceptual data model was analyzed and a logic design was achieved in multi-dimensional database mode. The optimal data model was confirmed by identifying data objects, defining relations and reviewing entities. The conversion of relations among entities to external keys and entities and physical attributes to tables and fields was interpreted completely. On this basis, a multi-dimensional database that reflects the management and analysis of a dam safety monitoring system on monitoring data information has been established, for which factual tables and dimensional tables have been designed. Finally, based on service design and user interface design, the dam safety monitoring system has been developed with Delphi as the development tool. This development project shows that the multi-dimensional database can simplify the development process and minimize hidden dangers in the database structure design. It is superior to other dam safety monitoring system development models and can provide a new research direction for system developers.

Key words: *dam safety*; *multi-dimensional database*; *conceptual data model*; *database mode*; *monitoring system*

DOI: 10.3882/j.issn.1674-2370.2008.03.012

1 Introduction

On the basis of the one machine and four bases principle, a dam safety monitoring system can give a quick and reasonable appraisement of the dam status in response to monitoring data, playing a vital role in real-time data analysis (Wu and Gu 1997; Wu 2003; Gu and Wu 2006). However, there is too much monitoring information stored in the system, and if the data model has been designed unreasonably, the system will generate redundancies, anomalies, inconsistencies and other problems even if a good database management software is used. These problems make it very difficult for the monitoring system to perform at its highest level, which means the pros and cons of the monitoring database design will directly affect system quality and operation (Xu et al. 2006). Therefore, an effective database model that improves system performance is urgently needed.

The multi-dimensional database model often uses data mining technology to determine the changing trend of monitoring data. The model can be effective for a large number of monitoring data summarized and analyzed in terms of data changing processes, changing trend

---

This work was supported by the National Natural Science Foundation of China (Grant No. 50539010, 50539110, 50579010, 50539030 and 50809025).
*Corresponding author (e-mail: *zefherohero@yahoo.com*)
Received Apr. 29, 2008; accepted Aug. 13, 2008

and changing amplitude. Furthermore, it can also be used to analyze reasons for varying trends and amplitudes of data and relations among those reasons. As opposed to traditional databases, the multi-dimensional database model can ensure that tables do not produce abnormalities and data redundancies and are able to support user-specified requests (He and Zhang 2006). Moreover, this data model allows for making changes in data and service requirements to make the database more accurate and database maintenance more effective. This paper analyzes the design process and implementation of the multi-dimensional database in a practical system development project.

2 Multi-dimensional database system development

The key techniques in the system can be shown through the development process of the multi-dimensional database system (Figure 1). The process shown on the left side of Figure 1 is the database structure design and the one shown on the right side is the database behavior design. Connections that work on both sides display a mutual influence. System developers should first construct an E/R (entity and relationship) model of monitoring information through data analysis. Then comes the key period in database structure design, database logic design (Xu et al. 2006), which plays a critical role in improving the efficiency of the database system.

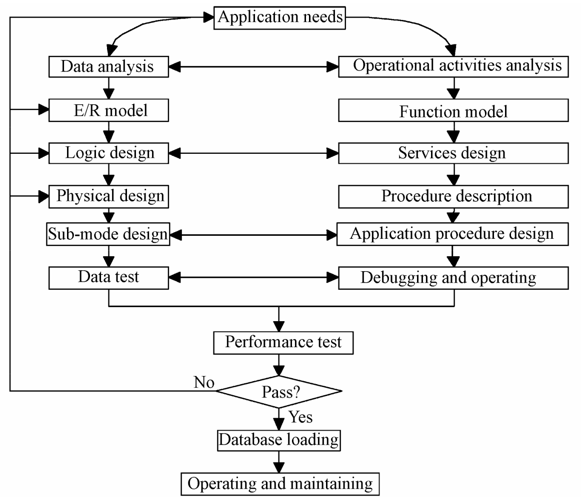

Figure 1 Development process of the multi-dimensional database system

As a special relational database, the multi-dimensional data model only requires data to be expressed as a relation form on the conceptual level and does not specify the way that data should be stored physically (Stöhr et al. 2000; Goil 1999). Basic tables and operation results are all expressed as relation forms on the conceptual level, which allows one operation result to serve as the input for another. Therefore, complex and public operations can be packaged and these processes can be reused when the need arises. Because the multi-dimensional conceptual data model remains the technical core of the database logic design, it should be analyzed first.

3 Multi-dimensional conceptual data model development

3.1 Identifying data objects

After the preparatory system development work is completed, a series of original documents should be collected or created, including report examples and input tables provided by reservoir personnel and work process documents designed according to regulations. When creating a data model, raw monitoring information should be reviewed first and all data information the system deals with should be listed next. Then, analysis can begin on the work process of the core project, which often contains the majority of entities.

In addition, most of the work process is triggered by data reports and analysis statements. After the data template is identified, all information contained in this work process needs to be written down. Later, entities or attributes among the data information need to be distinguished. At this time, attention should be paid to recording all repeated data, and those data that have been confirmed in the work process analysis but are still missing should be contained. When the list has been edited, entities, attributes and relations can be extracted. Each item of every list should be determined whether it is a target or facts in the target. Then, targets become entities and facts become attributes in turn.

3.2 Defining relations

A general understanding of analyzed entities and attributes should arise after reviewing all monitoring data documents. Then, relations among these entities should be established and attributes and constraints of every entity should be reviewed.

Although attributes can be checked first in theory, it is more effective to begin with their relations. That is because some projects in those monitoring data documents may become additional entities and others may be required to add attributes for identified entities.

Later, the definition of relations can begin with the core entities of the system. Then, other entities can be added and linked with them.

At last, relation types (one-to-one, one-to-many, many-to-many) can be defined in the process or lines can be painted simply to show that there is an existing link, which will be followed by more detailed analysis.

3.3 Reviewing entities

Each entity should undergo a detailed analysis when all entities and their relations have been obtained. The following information needs to be identified during this process: (1) Relations between entities and the problem domain; (2) The work process of creating, modifying, using and deleting entities; (3) Other entities that may interact with or depend on each other; (4) Business rules and constraints belonging to the entity; (5) Entity attributes.

3.4 Constructing the conceptual data model

The multi-dimensional conceptual data model includes descriptions of entities, attributes and their relations (Nguyen et al. 2000). It defines the data usage for the whole system, including not only the logical model, but also the interaction between the work process and monitoring data. The E/R map of monitoring point information is shown in Figure 2. For each relation pair in Figure 2, it is necessary to determine the relation base, the option feature of each participant, and any of the attributes and constraints of the relations.

The purpose of the monitoring system, main functional modules and working performance should be clear. It is also necessary to establish design standards for the monitoring system. Then the scope of the monitoring system should also be defined, as it can determine the functional modules to be included in the system.

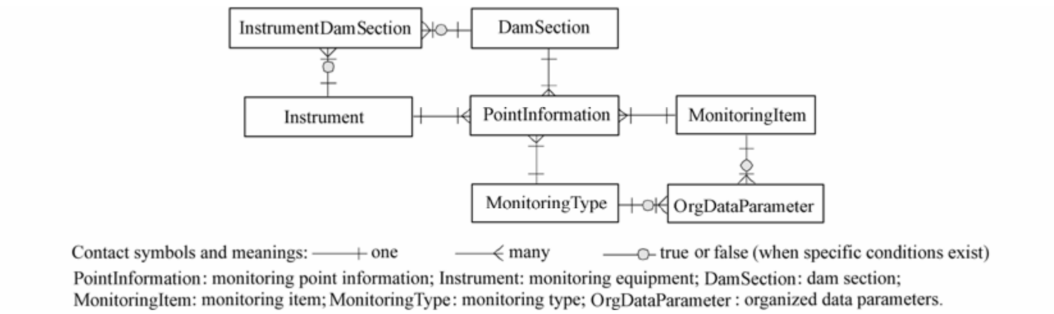

Figure 2 E/R map of monitoring point information

4 Logic design of multi-dimensional database mode

Multi-dimensional database mode means that the definition of the data physical layer is the realization of tables and views that change a conceptual model into some physical expressions that the database system can achieve (Kong 2006). It is another form of expression the data model describes from the database engine—for example, tables or triggers. The database mode of the monitoring system converts the conceptual data model into a physical form and describes the physical planning of tables, including a description of tables and the physical structure of the data.

4.1 Design process

Logic design changes the E/R model into an equivalent relational mode. Entities and some links can be expressed as tables while attributes in the E/R map can be converted into table attributes. Specifically, links among entities are converted into external keys, and entities and attributes are converted into tables and fields. The conversion process is as follows:

(1) Each entity in the E/R map is converted into a corresponding table, which should include all attributes of the corresponding entity. Keywords are determined based on the expression meaning of the table.

(2) If a 1 : 1 relation remains between two entities, the two entities are converted into two relational table modes. Then, their keywords are inserted into the other side. Examples include monitoring instrument (instrument code, item code, instrument parameters, etc.) and monitoring item (item code, instrument code, item name, etc.).

(3) If a 1: m relation remains between two entities, keywords from side “1” will be inserted into the corresponding table of side “m” as external keywords and attributes of relations will also be inserted into side “m” at the same time. Examples include dam section (section code, section name, etc.) and monitoring points (point code, section code, monitoring item, point name, etc.).

(4) If a 1: n relation remains in one entity, one more attribute should be created in the corresponding table to represent the keyword of the higher level linked to the entity. For example, the strain gauge has three-oriented, five-oriented, seven-oriented and nine-oriented types in an instrument table. Therefore, the attribute’s monitoring item number, monitoring direction and code can be set up to distinguish these instruments.

(5) If an m : n relation remains between two entities or within one entity, a table needs to be constructed separately to link these two entities or individualities. Attributes in the table should at least contain keywords of the two entities or individualities. If there are attributes that have existed in relations, such as monitoring instrument (instrument code, instrument name, instrument parameters, etc.), dam section (section code, section name, etc.), or dam section equipment (section code, instrument code, etc.), they should also be incorporated into this table.

In the database mode of the monitoring system, the definition of the table is derived directly from the conceptual data model. Entities become tables and each field in these tables corresponds to entity attributes. In the process, attention should be paid to constraints, relations and indexes. Restraints of most domains and attributes will become field-level constraints of the database mode, while restraints at the entity level are usually converted into table constraints. Furthermore, entity integrity constraints can often be achieved through the definition of a primary code for each table. In addition, the necessary framework of the monitoring system should be determined in database mode design. When using SQL Server to achieve the logic framework of a client/server application, the server will implement the data operation and the client will be responsible for the response to users. Then, the database mode will include each table, view, storage process and trigger of the database.

4.2 Preparing for multi-dimensional database mode

The structural mode of the multi-dimensional database based on the E/R model is shown in Figure 3 (incompletely). It is defined by a factual table and many dimensional tables, called star schema. Although not all attributes of the dimensional tables have been shown completely, dimensions do display completely. Generally speaking, the structure of the dimensional database is always similar to different systems. It always includes a factual table and many dimensional tables, though the numbers of dimensions, facts and attributes are different. However, each entity should be dependent on the primary key in the database mode.

4.2.1 Factual table

The factual table includes some primary dimensional yards and facts. The table is used to model monitoring data. The dimensional yards are key codes that link the factual table and dimensional tables, while the facts are monitoring data, representing the actual measured values. The appropriate size should always be used in table design (size here means details of the monitoring data and original data category). The bigger the size, the more analysis methods there are and the more effective the system is, but the more complex data analysis is. In addition, it should be noted that not all historical monitoring data can be taken at the same size level and special analysis is needed when putting senior facts into the junior level. Usually, the factual tables can be constructed according to the most junior level of the monitoring data.

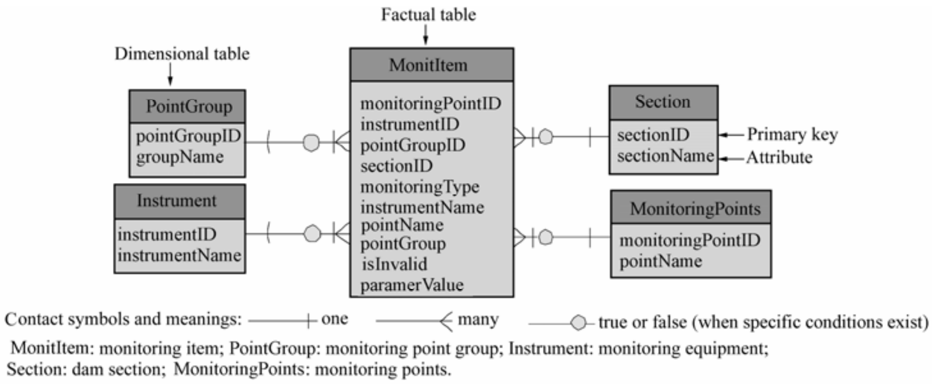

Figure 3 Structural mode of multi-dimensional database

There are service, snapshot and coverage tables in the system (He and Zhang 2006). The service table is the most common form used to model. The snapshot table can record measured values at a specific point in time and describe the status of a dam through periodic snapshots, which is a very effective way to draw value-varying maps. Snapshot and service tables are often used in conjunction. Coverage tables may not contain any attributes but include a line of data with special dimensional marks. These dimensional marks define the incident completely, and the data will show that the incident has already occurred.

4.2.2 Dimensional table

The main purpose of dimensional tables is to provide descriptions for attributes. Users can choose them as conditions to deal with data to appraise the dam status. In some cases, the dimensional table can illuminate the size of the factual table and redundant data contained in it, which is available for data mining. Tables are rarely changed in the multi-dimensional database, so the problem of updating a special data item in some positions doesn’t exist. The majority of the many-to-many linkages often exist between these dimensions, but when there are multiple contacts between the factual tables and dimensional tables, it is necessary to standardize these dimensions, which means snowflake schema should be used in this case (Jiang et al. 2005; Jensen et al. 2004). This situation is very rare, as it can be eliminated through adjustment of the size of the factual table.

4.3 Thoughts on logic model design

Attention needs to be paid to the following points when designing a dimensional database model: (1) Monitoring data should be original in the factual table. They should not be summarized. These data should also be at the same level and of an appropriate size. (2) The reference integrity of monitoring data should be ensured. All monitoring information existing in the factual table should emerge in the dimensional tables during design. (3) The star pattern model should be used as much as possible. When a snowflake pattern model is used, the number of dimensional tables linked to other dimensional tables needs to be minimized.

5 Application procedure design and implementation

Application procedure design includes service design and user interface design. Service design is the design of an interactive process between users and data to take the system into data analysis. It must be ensured that all features and required instructions are reflected in the database application procedure design. In addition, the interface represents the system to most users, so an appropriate and user-friendly interface for database applications is needed.

5.1 Service design

Service design refers to the management and analysis of dam monitoring information. These matters must be included in the monitoring database system to ensure that the database is consistent with reality and can support user information needs. Service design aims to define and incorporate high-level features that the monitoring system needs, including the data usage service, functional features of service, output service, the importance to users, and the expected rate of usage (Xu et al. 2006). In a dam safety monitoring system, there are three service types:

(1) The searching service searches monitoring data information that will be shown on the screen or remain the other branch input, such as the drawing time history plot, positional plot and correlation plot, considering each summarized eigenvalue and the modeling of monitoring values based on the statistical model, BP neural network model, time series model and gray model in order to forecast dam status.

(2) The updating service inserts a new record, deletes an old record, or modifies existing records of affairs in the database, including managing monitoring instruments, monitoring items, or adding, editing and deleting monitoring points. In addition, manual and automatic input of monitoring data, data integration and other operations can also be performed.

(3) The mixing service searches data and updates data services, including making data reports, searching and preparing for monitoring indicators, searching and inputting inspection information and managing inspection projects and dam safety documents.

5.2 User interface design

The interface is one of the most important components of the system. After necessary functions have been developed, the appropriate user interface should be designed for the database application. In the design of data reports from the monitoring system, attention needs to be paid to the following: visual form, report layout, meaningful titles, logical combinations and sequence of fields, consistency of terminology and acronyms, consistent color, visualization of space and boundaries used in data fields, error message alarms for unacceptable values, clarity of optional field markers, interpretative information for the field, and so on. The interface should be friendly in its display of the required information. If the interface is easy to master and intuitive to operate, users will find it easy to take full advantage of the information provided by the database.

In the design process of the user interface, the construction of the system interface must be established according to the work process supported by the system rather than on the basis of data structure. This means that the interface must focus on the tasks that the user performs and build the system to support these activities. There are two options in building the monitoring system: a single document interface system—using a single window to display data, and a multi-document interface system—using multiple windows to meet requirements. Specifically, it should be decided according to the work process what kind of data should be included in the form, and then the layout of the form and the option of certain control components should be determined depending on the expression structure of the actual data.

5.3 Implementation of dam safety monitoring database system

Based on the modeling principles of the multi-dimensional database described in the text, the dam safety monitoring system uses Delphi as the development tool. Compared with traditional database systems, the constructed multi-dimensional database model has shown a satisfactory performance in system development. For example, a rapid response of the multi-table inquiry has been realized in the system. A dam displacement positional plot containing data information from 36 monitoring point tables is shown in Figure 4.

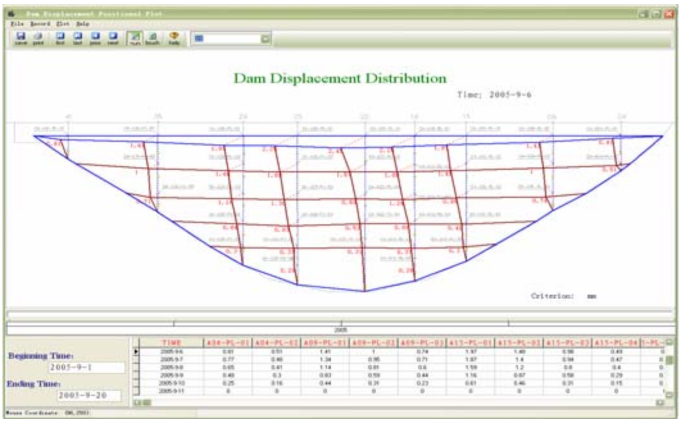

Figure 4 Dam displacement positional plot

The system interface takes the menu form. It can facilitate real-time data collection, processing, integration and analysis as well as appraisement of the dam’s safety status. It can provide significant help in expert decision-making analysis. Moreover, information visualization technology incorporated into the system can help in the building of a user-friendly interface to effectively manage dam monitoring information. In a word, the development and research of the system has become an important basis for the completion of the project and will provide a security guarantee for dam operation in the future.

6 Conclusions

(1) In order to improve the effectiveness of dam safety monitoring systems, a multi-dimensional conceptual data model has been built and the logic design of the multi-dimensional database mode has been analyzed in depth. The optimal data model has been constructed based on dam safety monitoring information. The multi-dimensional database system is not only able to access monitoring data effectively to meet user requirements, but can also be accepted in current computer hardware and software environments.

(2) Service design of dam safety monitoring systems has been divided into the searching service, the updating service, and the mixing service, in order to analyze the system comprehensively. The practical requirements for user interface design have been listed and the foundation for construction of the system form has been explored in depth.

(3) The practical development project shows that the multi-dimensional database can minimize the hidden dangers in database structure design and simplify the development process. Moreover, it can strengthen the system stability. The effectiveness shown in the dam safety monitoring system proves that the multi-dimensional model is superior to other models and will generate a new direction for dam safety monitoring system developers.

References

Goil, S. 1999. *High-Performance On-Line Analytical Processing and Data Mining on Parallel Computers*. Ph.D. Dissertation. Chicago: Northwestern University.

Gu, C. S., and Wu, Z. R. 2006. *Safety Monitoring of Dams and Dam Foundations: Theories and Methods and Their Application*. Nanjing: Hohai University Press. (in Chinese)

He, Y. J., and Zhang, J. C. (trans.) 2006. *Designing Effective Database Systems*. Beijing: China Machine Press. (in Chinese)

Jiang, W. W., Xiong, D. P., and Zhang, X. X. 2005. Research of MOLAP storage and query techniques based on multidimensional database. *Computer and Digital Engineering*, 33(2), 56-59. (in Chinese)

Jensen, M. R., Holmgren, T., and Pedersen, T. B. 2004. Discovering multidimensional structure in relational data. *Lecture Notes in Computer Science*, 3181, 138-148.

Kong, D. S. 2006. *Research on the Technology of Multi-Dimensional Data Storage*. Ph.D. Dissertation. Wuhan: Wuhan University of Technology. (in Chinese)

Nguyen, T. B., Tjoa, A. M., and Wagner, R. R. 2000. An object oriented multidimensional data model for OLAP. *Lecture Notes in Computer Science*, 1846, 69-82.

Stöhr, T., Märtens, H., and Rahm, E. 2000. Multi-dimensional database allocation for parallel data warehouses. *Proc. 26th International Conference on Very Large Data Base (VLDB)*, 273-284. Cairo: IEEE.

Wu, Z. R., and Gu, C. S. 1997. *Dam Safety Comprehensive Appraisal Expert System*. Beijing: Beijing Science & Technology Press. (in Chinese)

Wu, Z. R. 2003. *Safety Monitoring Theory and Its Application of Hydraulic Structures*. Beijing: Higher Education Press. (in Chinese)

Xu, L. F., Peng, B., and Wu, Y. Y. 2006. *Database Design and Realization*. Shanghai: Shanghai Jiao Tong University Press. (in Chinese)