Data Warehouse Architecture for Metal Mine Enterprise with UML AND CWM

Xinrui Liu [1, a]， Guoping Lei[1,b]， Fengyu Ren [2,c] and Hongbin Ma [2,d]

[1] Institute of land management，Northeastern University，Shenyang, China

[2] School of Resources and Civil Engineering, Northeastern University,Shenyang, China

[a]neu-lxr@163.com, [b] guopinglei@126.com, [c] happyrst@163.com, [d] neu_mhb@163.com

Keywords: data warehouse; metal mine enterprise; object-oriented technique; CWM; multidimensional conceptual model; data analysis.

Abstract. Owing to the curse of inflated data and deficient information knowledge, we presented a new approach to construct data warehouse architecture for supporting data analysis and decision making with object-oriented and CWM (Common Warehouse Metamodel) technologies based on comprehensive analysis of current situation of metal mine enterprises in our country. Firstly, basic data of metal mine enterprises were analyzed in distribution from data statistics and production operation two points of view. The former are targets for data mining arithmetic; the latter are the members of data marts built by distinct duty of department. Then, in the light of strategies of object-oriented data warehouse, it is reasonable to go through three steps of source data area, the staging area and the presentation area under a unified spatiotemporal and dimensional framework. According to this, new dataflow architecture was put forward with CWM to implement both transaction database for metal mine daily management and data warehouse for decision making. Obeying unified modeling language (UML), we built concept model of metal mine entities from real world to information world, which was subsequently mapped into dimensional model of data warehouse correspondingly depending on dimensional mapping rules. It shows that: for metal mine having tremendous multidimensional mineral entities, it displays more technical advantages of data warehouse of smart mine with object-oriented technique and CWM than structuring technique , so is for data integration.

Introduction

In recent decades, especially due to rapid development of geospatial information, the digital and informational production is emphasized by more mine enterprises. However, there are different conditions and digital degree among mines in China: some extensive mine enterprises may have excellent operation including independent digital products, but others may not have all or even not any [1]. On the other hand, because of poor compatible MGIS systems and multiple storage styles of data in various carriers, the architecture of data integration platform are extremely confused within the mine enterprises without a unified planning and structure among transaction-level databases, central data warehouse and data marts of departments, which makes mine face the awkward situation of magnanimity, jumbled, changeable data accumulated and seriously deficient information knowledge and does the data query and decision-making for mine companies harder. Therefore, it is a critical problem how to steer these data and turn them to valuable information, avoiding information islands. Considering that "smart" [2] usher in the brand new trend of building information system, object-oriented data warehouse with CWM, as data scientific organization and management technology for decision support system [3], is not only one of key technologies of digital mine, but an effective approach to discover knowledge intelligently from magnitude data and end up solving the issues of data explosion and information island as well.

According to Joseph [4], Object-Oriented (O-O) techniques are much more powerful for modeling metal mine enterprise than structured techniques (such as ER model, enhanced-ER model), because they reflect the policies that business people want to use in running the enterprise and building systems that implement these policies. Metallic mineral resources, constituting mineral resource with non metallic mineral resources and energy resource, becomes deadly difficult points of digital mine because

of the features [5], such as complex geological conditions , changeable mining methods and uncertain accidents, especially underground metal mine covering from the surface to the underground and contain multiple subjects of data. For these reasons, it is more beneficial to unfold multidimensional spatial entities of metal mine through measures of building data warehouse with object-oriented technique.

In this paper, we research and analyze the approach for constructing an object-oriented data warehouse of underground metal mine data. In this technique, after the illustration with respect to the ontology of mine dataset and business themes in section 2, the framework of data warehouse is primarily discussed for metal mine information system in section 3. We subsequently express the first abstract concept of mine entities from real world to information world with unified modeling language (UML) schemas. What's more, on the basis of schema transformation rules, relation schemas are mapped into the fact-dimension tables of data warehouse correspondingly. These are main content of section 4. Section 5 presents the conclusion.

Metal mine enterprise data and research

Analysis of metal mine enterprise data resource. Metal mine data resource is of abundant content, covering a wide range, various data styles. It is grouped into four parts including production data, office data and Enterprise external data [6].
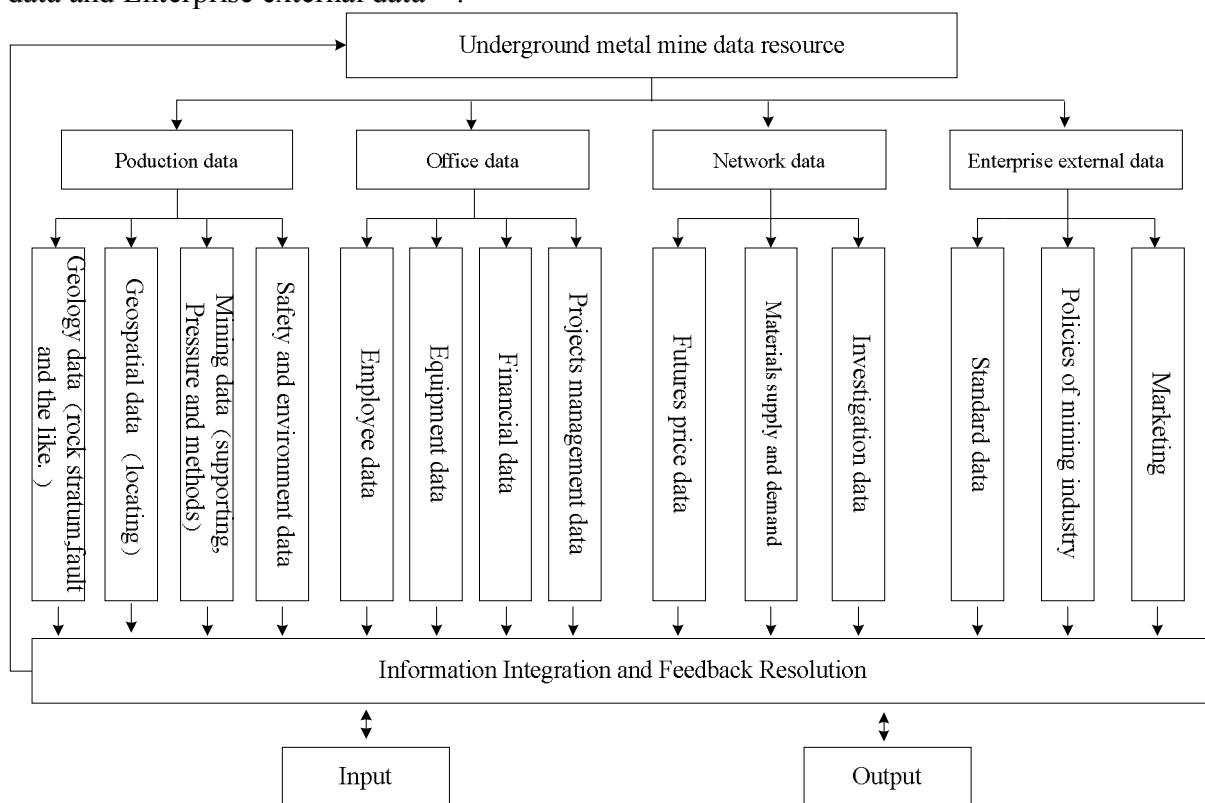

Fig.1 An flow design of metal mine enterprise data resource

Production data can call business datum mean those collect from the mine and subordinate enterprise's existing business process system to keep in business process system and Daily basic key data of mine of producing and managing one grade of data of relevant affairs, for instance geological prospecting, mine measurement, mining methods, pressure control, supporting measures and ore quality control ,etc. From the view of production process, they cover from real time current production data to history accumulated data. From the point of subjects, they include mining survey, geology, mining engineering, safety and environment engineering. This part of data source is core of a productivity effect of a mine enterprise and mine basis of exploiting the craft.

Office data are the office system data inside company, factories and mines to work, worker's information, apparatus information and financial data, etc. work information data, these data are shown

as electronic data and non- electronic data in form. The former are saved as the styles of spreadsheet, database and text processing document and so on; the latter consist of a great number of official documents, such as files, notification and meeting summary.

The network data are enterprises through Internet data obtained, these data can be obtained through the e-commerce system of enterprises, can investigate and obtain through the network too. There are two parts: one is common business information including the spot price and forward price of metallic product given by metallic product and their supply and demand information monthly given by State Statistics Bureau in our country and so on. Two is transmittal data according to certain standards, usually consisting of techniques in batch for data changing and trading.

Enterprise external data refer to operate, have, datum controlled by enterprise those, these data are full of electronic forms, for instance research paper of the consultancy service of the marketed.; Some are non- electronic forms, for instance policy related to mining industry, regulation file which the country issues and so forth .

Turning Mine Data into Objects. Considering the comprehensive and unpredictable metal mine data, it is primary thing to research the type of data and the attributes.

Raw data can be represented categorical, ordinal and quantitative variable at statistics or data mining standpoint [7]. Data of underground mine are appropriate for different types of data and include all the variables. Data from Mine Enterprise whose goal is to pursue profit never lack quantitative variable. According to operation, we can divide them into six datasets.

Geospatial location: the sites of factory buildings depend on the places of ore, which affect by lives, traffic, power supply, water supply and mineral supply of local place and other situations as well. The information keeps changing and accumulating in the process of mining development.

Geological conditions: The value of mineral resources and mining conditions impacts the economic benefits of Mine Enterprise. Many factors, such as value of ore, amount of reserve, size and occurrence of ore body, burial depth, the stability of ore and the distinct measures of mining, affect the investment of construction and cost of mining.

Mining technology: the uncertainty and unpredictability of geological conditions and constructions, geometry and chemical and physical properties, these factors cause mine face to constantly change and move. Simultaneously, methods, technology and parameters of mining update along with them. Therefore, the information of mining technology continuously alters in order to enhance the management and adjust production schedule and excavating capability.

Mechanical equipment and ventilation: the process of production for mine has multiple steps including blasting, loading, transport, unloading and ore processing and so on. Assistant things for mining are applied, such as machines, electric power, drainage and ventilation. How to control and steer them will produce different economic benefits.

Safety and environment: there are large amount of dust working in mine, especially underground. All these mean harsh operating surroundings and low security for workers. Mineral resources are generally non-renewable; we must take care of them and excavating scientifically for our offspring. So we must pay serious attention to safe information and environmental data at every time.

Finance and markets: Some financial information and markets situations containing operating costs, sales of ore and domestic and international policies and markets will have some effect on decision-making for manufactures and economic interests.

Compared two above opinions, the fact is found that the former are used for data analysis and the latter are more appropriate for follow-up construction of data marts with UML.

Strategies for data warehouse of Metal mine enterprise information.

It is convenient and economized for most mine enterprises to build and maintain a unified aggregate data warehouse because of fast querying and transformation. However, it may be other way round for a grouped company owning multiple subsidiaries. Just like Ralph Kimball[8] says there are no two "contrasting" points of view about top-down vs. Bottom-up data warehouse, at least the techniques' boundary of building data warehouse are not very strict and clear. Neither approach taken to these limits is unfeasible and unsuccessful. The better workable solution is a blend of the two approaches

where we put in place a proper and uniform space-time architecture that guides the design of all the separate pieces, successively noticing and improving according to the feedback.

After arguments about mine data，the fundamental database is firstly constructed with Object-Oriented Software Engineering (OOSE) technique, whose objective is not only to optimize the quality of data by partly cleansing and standardizing but also cater to the requests of transformation level information management system. Next, data extracting from database is pretreated and integrated through ETL toolkit. The object classes are mapped to fact-dimension model by the rules. Eventually, the data marts are separated from central data warehouse depending on duties of departments for decision-making.

Inside information of metal mine enterprise should be able to communicate each other because of association between businesses of metal mine enterprise. Metadata are data about data. Inclusive data structure and semantics in metadata is important evidence for data integration. Meanwhile, It creates a query and update bridge between heterogeneous information. Therefore, according to Common Warehouse Metamodel[9], Central Meta Data Repository is built for data warehouse integration strategy, removing the communication obstruction among data marts. Figure 2 shows the entire framework.
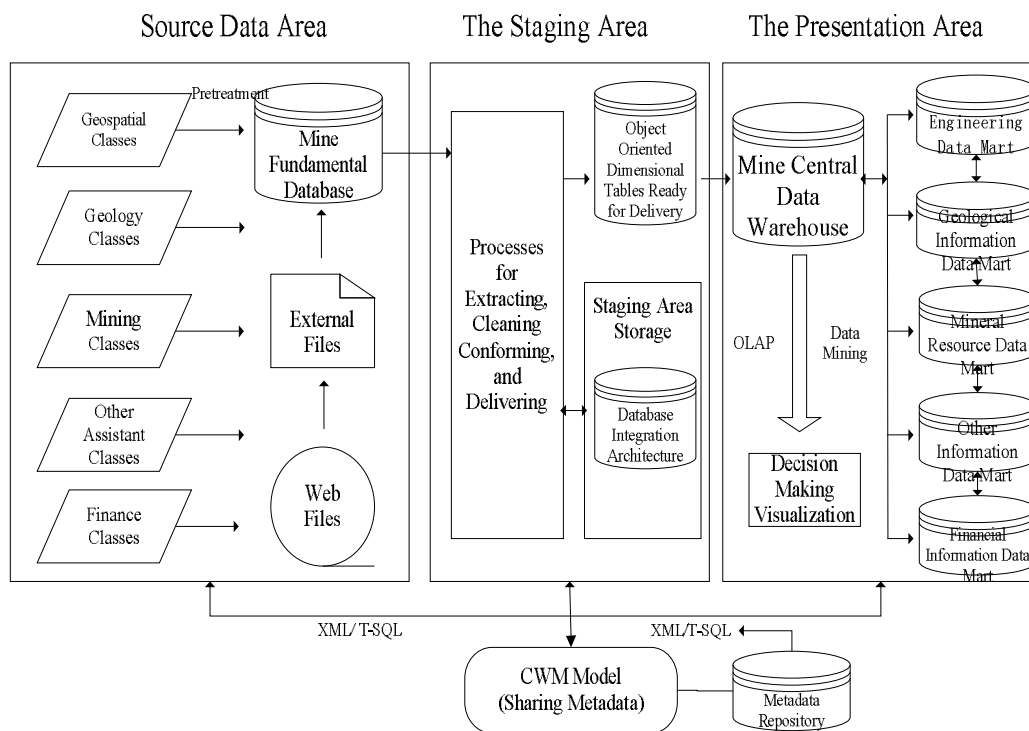

Fig.2 Architecture of object-oriented data warehouse for metal mineenterprise

Mine enterprise data warehouse building with object-oriented dimensional model

Themes of mine data warehouse designing. The raw data of metal mine production contain the transaction level data that has been stored in transaction processing system and being produced by mine daily production operations, including the engineering entities of mining area, the geological entities of mining area, the mining design and technology entities of mining area and other aided mining projects' entities of mining area:

Engineering information management: As the geospatial information basis of subsequent themes adding safety information supervision and environmental protection information, we can always control the process of building engineering projects to realize the green and safety production scheduling and management. There some temporary or permanent project entities, such as tunnel engineering projects, and other artificial projects (the bolts or anchor cable for construction safety).

Geological information management: It is used to analyze three dimensional geological structures and provide decision information. Its original data are geological entities of mining area and part of engineering entities as the most important and difficult three-dimension spatial entities to express.

Mineral resource information management: It is applied to analyze regularity of mineral resource and conditions of ore removal and evaluation of the effects of mine methods. The information source from the process of engineering projects and geological jobs gradually displayed including mining methods and flow and so on. Its main part is of some transaction data of mining design and technology entities, involving some geological entities and engineering entities as well.

Other equipments and materials information management: The aim is to design optimum plans for application, scheduling and improvement of equipment management. They refer to all data resources except the data which have been discussed above in huge mine system, which are mainly from other aided mining projects' entities of mining area and also contain a small part of engineering entities.

Mapping rules between dimensional conceptual model and UML. Measure or a dimensional model is typically organized around a central theme. This theme is represented by a fact table. Facts are numerical measures and principal statistical objects [10]. In contrast, use case of O-O technique is related to theme fact that bridges a cohesive tie between external and internal system. We adopt the plan that one use case picture represents one theme. Due to the research of themes and strategies from above paragraphs, we map use case diagrams of UML to stars schemas and fact constellation schemas for the whole multidimensional conceptual model in order to implement the construction of complex and enormous data warehouse. UML classes are mapped to fact tables and dimension tables and the relations of various classes are literally similar expression to those among fact tables and dimension tables [11]. The granularity of use cases [12] decides the numbers and categories of attributes of model tables.

Table1 Mapping relations between UML and multidimensional data model

UML Objects	Fact Dimensional Model Objects
Use cases	Themes
Granularity of use cases	Numbers of dimensions
Classes	Fact tables or dimension tables
Attributes of classes	Fields of fact tables or dimension tables
Class relations	Relations of fact tables and dimension tables
Class methods	Operations for multidimensional data

Implementing mine data warehouse. According to requirement analysis and theme information of mine business process, class diagram is built in figure 3.
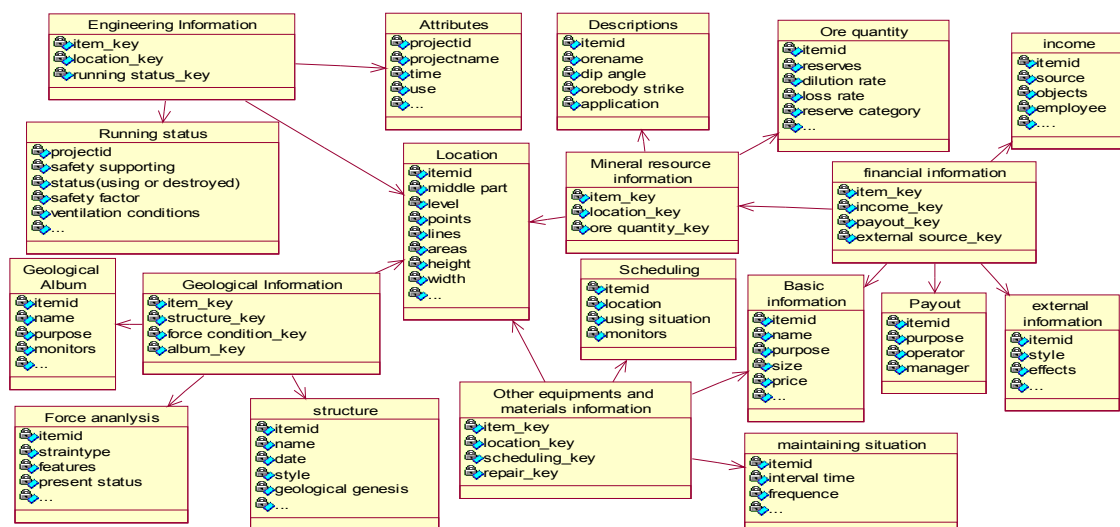

Fig.3 The class pictures of the whole metal mine information

Conclusions

To strengthen scientific management and avoid the fatal mining mishaps, we need to pay more attention to research and develop methods and tools of data analysis such as accident warning or data mining. The object-oriented theory and the approach for constructing data warehouse and metadata model with CWM should be both taken into account in this topic. Following conclusions can be drawn from this case study: Owe to multidimensional and complex features of metal mine data, the encapsulation, inheritance, polymorphism of object-oriented skill may demonstrate the more accurate associations and discrepancies between data objects, simultaneously reducing more redundant. As for CWM, it is more benefit to data sharing. Because of the time limitation of mining development, metal mine enterprise constructs not only the transaction-level database for satisfying the management of daily operation, but also builds data warehouse for OLAP and data mining in order to effectively control and monitor data by all levels of users.

Acknowledgements

References

[1] Hongbin Ma, Xinrui Liu, and Hongdi Zhao,et al, "Study on the construction of metal production information management system with UML," International Conference of Environmental Science and Information Application Technology (ESIAT), 2010 , U.S. IEEE Press, July, 2010, pp.614-615.

[2] Daqun Qian, Smart Earth in China, IBM, 2009

[3] W.H.Inmon, Building the Data Warehouse (4th.ed.), New York:John Wileyand Sons Inc, 2005..

[4] Joseph M. Firestone, Object-Oriented Data Warehousing, Executive Information Systems, Inc.: White Paper No. Five, August 7, 1997.

[5] M.CxLipsett, CxR. Baiden, "Mining Information System Development," CIM Bulletin, vol. 94, 2001, pp. 63-66.

[6] Xinrui Liu, Hongbin Ma, and Hongdi Zhao,et al, "Research of Building Production-Oriented Data Mart for Mine Companies Based on Data Mining," 2010 Sixth International Conference on Natural Computation, IEEE Press, August 2010, p.2186.

[7] Jessica M. Utts, Robert F. Heckard, Mind on Statistics (3rd.ed.), USA: Duxbury Press, 2006.

[8] Ralph Kimball, Margy Ross, Warren Thornthwaite,et al., The Data Warehouse Lifecycle Toolk(2nd.ed.), USA:Wiely, 2008.

[9] John Poole, David Mellor. Common Warehouse Metamodel: An Introduction to the Standard for Data Warehouse Integration, USA,New York: John Wiley & Sons, Inc.

[10] Claudia Imhoff, Nicholas Galemmo, Jonathan G.Geiger, Mastering Data WareHouse Design Relational and Dimensional Techniques, Peking: China Machine Press, 2004, p. 19.

[11] S.K. Gandhi, Shailendra jain, "Data Warehouse Design: an Object Oriented Appraoch," Journal of Engineering, Science and Management Education, Vol.4, March, 2011, pp.16-17.

[12] Grady Booch,Robert A. Maksimchuk,Michael W. Engle,et al., Object-Oriented Analysis and Design with Applications(3rd.ed.), USA:Addison-Wesley Professional, 2007.