TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 57/67 pp354-361
Volume 13, Number S1, October 2008

A Decision Support System for Construction Projects Based on Standardized Exchanged Documents*

MA Zhiliang (马智亮)**, LU Ning (陆 宁), GU Weihua (顾卫华)†

Department of Civil Engineering, Tsinghua University, Beijing 100084, China;
† Department of Civil Engineering, University of California, Berkeley, CA 94720, USA

Abstract: Large amounts of documents are exchanged during the construction phase of projects, which covers the important management information. To utilize the exchanged documents to support decision-making of the management staffs, the requirement analysis was carried out based on the interviews to the practitioners. A decision support system called Explyzer+ was developed based on the previous prototype system Explyzer. The latter was enhanced by adding the functions to automate the whole process and the techniques of data mining including decision tree analysis and clustering analysis. A case study for in-depth information analysis was conducted based on the data obtained from a large construction project to demonstrate its feasibility and effectiveness. The system could effectively assist the management staffs to carry out in-depth information analysis for decision-making in construction projects.

Key words: data warehouse; on-line analysis processing (OLAP); data mining; decision support

Introduction

A construction project typically involves many participants such as the owner, the contractors, and the engineers and their coordination is normally carried out through large amounts of information that are exchanged in forms of documents, shop-drawings, and photos, etc. Among the documents, either in paper or electronic form, are not only the carrier of important management information for each participant, but also the formal evidence of the exchanged information between them or between the participants and the relevant government agencies. Usually more than ten thousand documents are exchanged in a large construction project, which are mainly submitted by contractors (including the general contractor and subcontractors), and reviewed and validated by the general contractor, the engineers, the owner, or the government agencies etc. to ensure the smooth progress of the project.

Obviously, the requisites for utilizing the exchanged documents are that the information in the documents is electronically stored in a structured form, and that a system facilitating the in-depth information analysis is available to the management staffs. Regarding the former, the exchanged documents could be generated and stored using XML[1], especially for simple documents that exist in a tabular form and filled with numbers, dates, or simple text. In addition, Caldas et al. used text mining approach to classify complex documents such as meeting minutes based on IFC objects[2]. However, this research was focused on analyzing the information in the simple documents, which constitutes the majority of the exchanged documents. The authors proposed a draft standard for the information exchange in construction projects based on XML and several fundamental standards for

Received: 2008-05-30
* Supported by the National Technological Promotion Program for the 10th-Five-Year Plan of China (No. 2004BA209B04)
** To whom correspondence should be addressed.
E-mail: mazl@tsinghua.edu.cn; Tel: 86-10-62773543

information exchange including ISO11179, CCTS, and UBL[3-5]. Regarding the latter, studies related to decision support system based on the information from construction projects have been conducted in recent years. Zhang et al. developed a prototype system using data warehouse to support the strategic decision-making of contractors[6]. Thammasak et al. established a data warehouse structure to be filled by extracting information from the contractors' databases, which may support the decision-making of construction firms[7]. Soibelman et al. applied decision tree algorithms and artificial neural network to analyze the reason for progress delay of construction projects based on the information accumulated in the databases[8].

However, to utilize the systems developed in these researches, the user has to not only spend a lot of time preparing the relevant input data, but also master some general-purpose tools for data analysis in order to carry out in-depth information analysis. The difficulties encountered in doing so seriously affect the enthusiasm of the management staffs to support decision-making by utilizing the information. Therefore dedicated systems that facilitate in-depth information analysis to the management staffs are required.

A research has been carried out by the authors in order to utilize the exchanged documents in construction projects, and a prototype system called Explyzer (named from "explorer" and "analyzer") has been developed[9]. However, the system could not be directly utilized by the management staffs in construction projects，since it just focused on the method of implementation. This research aims to provide a further solution for the utilization of the exchanged documents in construction projects by expanding Explyzer. This paper will shortly review Explyzer at first place and then present the new system Explyzer+. Finally a case study will be given showing how the system is utilized in a large construction project to demonstrate its feasibility and effectiveness.

1 Review of Explyzer

In another research, the authors developed a system called EPIMS (electronic project information management system) to support the collaboration among participants in the construction phase of projects, providing the functions for submitting, reviewing, and validating the exchanged documents[1]. As a result of using the system, the exchanged documents were accumulated in a structured form in the system and the system also facilitated the documents to be exported in form of XML files according to the draft standard for information exchange proposed by the authors as above-mentioned.

Explyzer was so developed that the standardized exchanged documents obtained from EPIMS could be utilized. The system consisted of two categories of functions, i.e., one for documents management and the other for in-depth information analysis. By using the functions provided in the system, the user could import the standardized exchanged documents into the data warehouse established in the system, and then carried out OLAP (on-line analysis processing) analysis by using the data warehouse and the data cubes in it.

Although the system demonstrated an approach to utilize the exchanged documents, two critical problems remained to be solved. One is that the process of information analysis is too complicated for the management staffs in construction projects. It is because the system has not provided sufficient functions for carrying out the in-depth information analysis so that the user has to directly use the general-purpose functions provided by the database management system. The other is that further methods for in-depth information analysis have not been implemented yet. The system supported only OLAP function, which left a great distance to achieve the goal of in-depth information analysis to support decision-making.

2 Development of Explyzer+

2.1 Requirement analysis

In order to discuss the direction of the further development of Explyzer, the authors showed some case study obtained from using Explyzer to the practitioners in construction projects. While they considered that the system was helpful, they gave some suggestions for improving the system in two aspects as follows:

(1) The system should cover and automate the whole process of the in-depth information analysis. Namely, the processes of importing the exchanged documents, establishing the data models, and analyzing the information should be fully supported by the system to prevent the management staffs from the tedious data

preparation work.

(2) Various data mining techniques should be implemented in the system in order to enhance the capability of the system for decision-making support. Meanwhile, the feasibility and effectiveness of the techniques should be demonstrated.

2.2 Design and implementation

Based on the above analysis, the authors developed the system model of Explyzer+ as shown in Fig. 1[10]. In this model, the standardized exchanged documents (simply called standard documents hereafter) could be imported from other systems and be filled into the data warehouse to establish the data cubes and data mining models. The whole process can be done automatically by the system after the user inputs the necessary parameters. In addition to the function for importing standard documents, two groups of functions were provided in the system. One group was denoted as the data cube tools and the other as the decision support tools. By using the former, the user could create, modify, and update the data cubes through user-guided interfaces. By using the latter, the user could not only carry out OLAP analysis, but also make use of data mining techniques, including decision tree analysis and clustering analysis, to support the decision-making for the management staffs in construction projects.

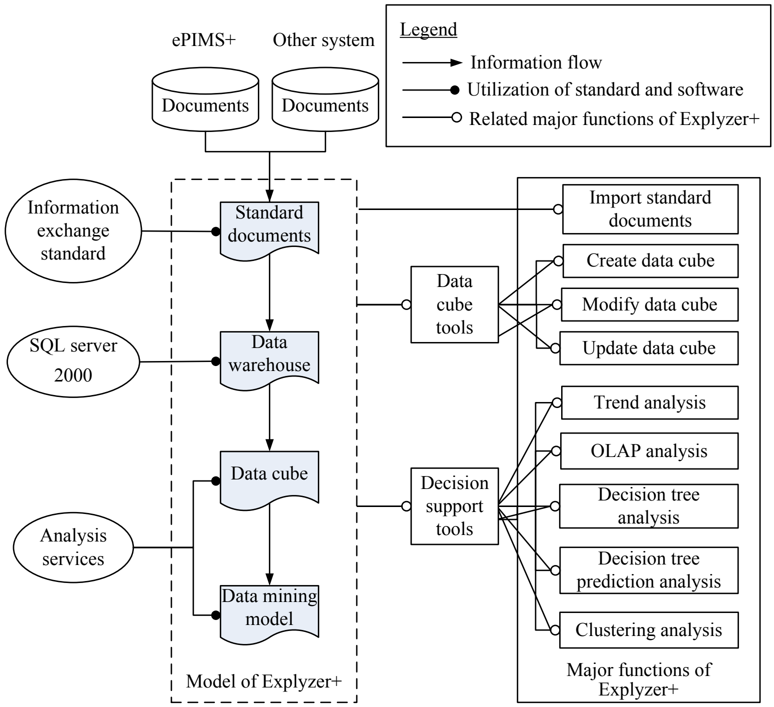

Fig. 1 Model of Explyzer+

Explyzer+ was implemented by adopting the B/S (browser/server) architecture. ASP (active server pages), VBScript, and JavaScript were used as the programming environment, and the Microsoft SQL server with analysis services was used as the database management system in it. The main interface of the system is shown in Fig. 2, where the user could click the function links on the left menu to enter a series of user-guided interfaces to go through the processes or click the buttons listed below the major area of the interface to execute corresponding commands.

Besides, in the implementation of the system, the DSO (decision support objects) component provided in analysis services was used to create, modify, and update data cubes and data mining models. Two ActiveX controls, i.e., the PivotTable in OWC (office web component) and the VSFlexGrid were used to browse and operate data cubes and data mining models.

3 Case Study

After finishing a certain part of projects, the general contractor usually wants to get an overall view of the

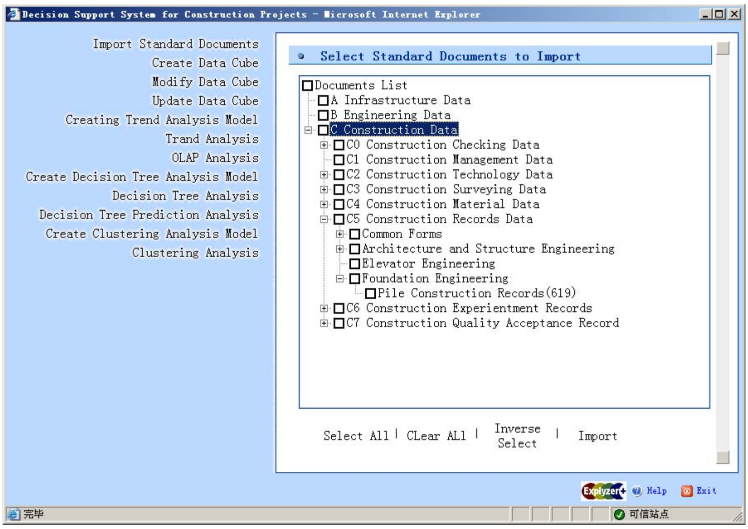

Fig. 2 Interface of Explyzer+

work and know which contractor's work is better. At the same time, other contractors may want to know if the quality of the construction is well controlled, and if not, what the influential factors are. These managerial intentions can be fulfilled by using Explyzer+.

Based on EPIMS, the authors developed the system ePIMS+ and applied it in the construction projects of the National Stadium i.e., the main stadium for the Beijing 2008 Olympic Games. Furthermore, standard documents were exported from the system[11]. In ePIMS+, by supporting filling the form on web, the system automatically generates the tagged files, i.e., XML files without introducing inconvenience to the user. The forms were displayed based on the document templates which can be established easily for various forms by using the system developed preciously[12]. By doing so, the authors have obtained a lot of exchanged documents of the project.

In the study, by analyzing the contents of the exchanged documents and carrying out numerous discussions with project participants, the exchanged documents were divided into four important categories, i.e., payment, schedule, quality control, and resource. Then 13 subjects in the categories were identified as having the potential to give meaningful result if they are analyzed[10]. Among them, those related to the quality control of pouring piles of the foundation were used as a case study for the in-depth information analysis.

Hence the exchanged documents on the quality control of pouring piles were processed by using the system to carry out in-depth information analysis as demonstrated in the following.

Step 1 Import standard documents and create data cube

"Import Standard Documents" function was used to import the standard documents that were exported from ePIMS+ into Explyzer+ at first. Then "create data cube" function was used to automatically create the data warehouse and the data cube corresponding to the subject "quality control of pouring piles" as shown in Fig. 3, through a series of user-guided interfaces. The information on each dimension is listed in Table 1.

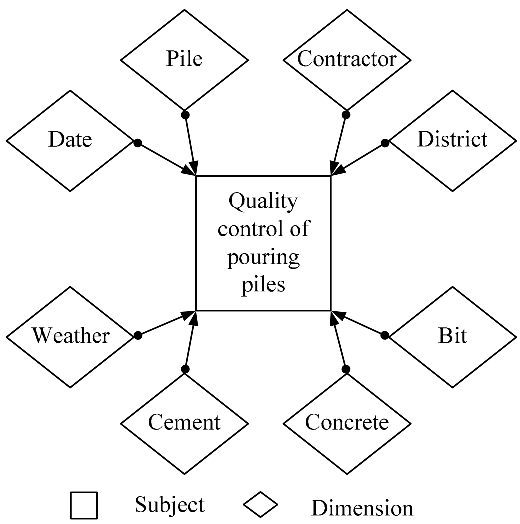

Fig. 3 Scheme of "quality control of pouring piles"

The major metrics "extra-pouring height" was focused in the case study. It equals to the actual pile elevation minus the designed pile elevation, which should

Table 1 Further information in scheme

Dimension	Information
Pile	Type, Diameter
Contractor	Contractor name
Date	Year, Quarter, Month, Day
District	District name
Weather	Weather keywords
Bit	Type, Diameter
Cement	Grade
Concrete	Grade

be in the range of 0.5-1.0 m according to the relevant standards to ensure the density of the pouring pile. If it is greater than 1.0 m, it means that the waste of concrete has occurred. Another metrics “filling coefficient” equals to the actual volume of concrete divided by the calculated volume of concrete, which should be in the range of 1.0-1.05 m according to the relevant standards, also to ensure the density of concrete. And it means that waste of concrete has occurred if it is greater than 1.05 m.

Step 2 Analysis with OLAP

The “OLAP analysis” function was used to observe the data cube, as shown in Fig. 4. With the interface, the user could choose different combinations of rows and columns such as “date + contractor” and “region + contractor” etc. to get a multi-angle observation on appropriate objects in the PivotTable.

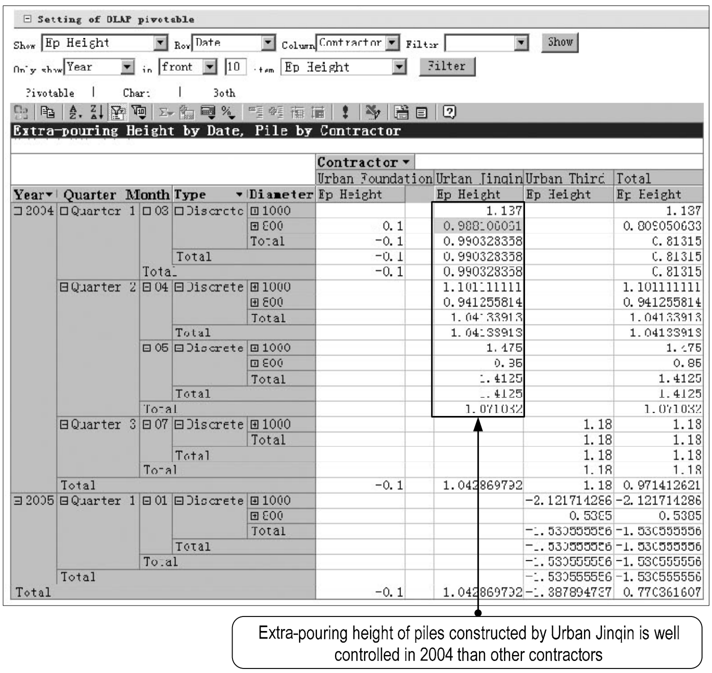

Fig. 4 Interface of OLAP analysis

In this case, “date + contractor” combination was selected and the average extra-pouring height was shown against them. Then pile dimension was added by selecting from the right-button menu of the date dimension to show the extra-pouring height under both date and pile diameter. From the result, it is known that one contractor named Urabn Jinqin controlled the extra-pouring height within range of 0.9-1.0 m in 2004, which is the best in the three contractors. It is also known from the figure that the average extra-pouring height of the piles with diameter of 800 mm is better (in the best region) than that of piles with diameter of 1000 mm (usually greater than 1.0 m) constructed by Urabn Jinqin. The former conclusion is useful for the general contractor to evaluate the subcontractor while the latter is useful for the subcontractor to find where the problems are in order to improve the quality control.

Step 3 Analysis with decision tree

“Decision tree analysis” function was used based on the established data cube in Step 1. The extra-pouring height was selected as predictive value, i.e., the observed or predicted value in the decision tree model.

While "contractor", "district", and "date" were selected as input value, i.e., the value corresponding the factors that affect predictive value in the model. The decision tree model thus established is shown in Fig. 5.

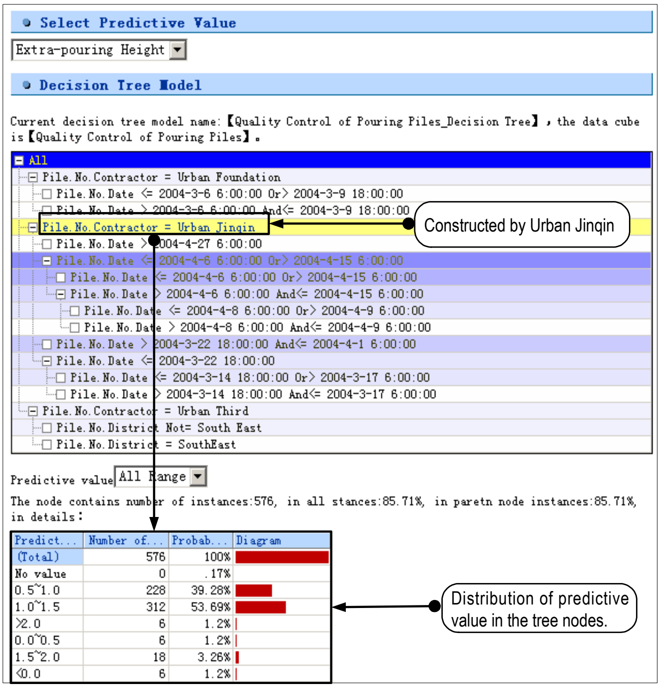

Fig. 5 Interface of decision tree analysis

A decision tree is a tree-based knowledge representation method used to represent classification rules showing the influence of each input value on predictive value. If one input value has more significant effects on the predictive value, it will be located closer to the root node of the tree model. From this rule, the major factors could be distinguished from the minor factors. Meanwhile, the number of instances of each input value in different tree nodes could be known.

From the figure, it is known that the most significant factor for the extra-pouring height is the contractor, because it goes on the root level of the decision tree. Then each node of the model can be clicked to find the corresponding distribution of extra-pouring height and to distinguish between the alternatives.

For example, Urban Jinqin constructed 576 piles, whose probability of controlling the extra-pouring height within 0.5-1.5 m is more than 93%. However the results of the other contractors are worse than that of Urban Jinqin. From general contractor's point of view, it means that the most important factor to ensure the construction quality is to select a good contractor. Since Urban Jinqin is proven to have worked well, if the future project's condition permits, the contractor should be employed to ensure the construction quality.

Step 4 Analysis with clustering

Based on the data cube established in Step 1, "clustering analysis" function was used, where extra-pouring height, filling coefficient, contractor, and district were selected as input values to set up the clustering model. Clustering algorithm is normally used to divide the object into several categories, and the objects have a high degree of similarity within each category and dissimilarity between categories. After categorization, the distributions of each input value can be examined to find the meaningful and interesting category. In addition, the result of clustering analysis could be analyzed by other data mining methods to get further results.

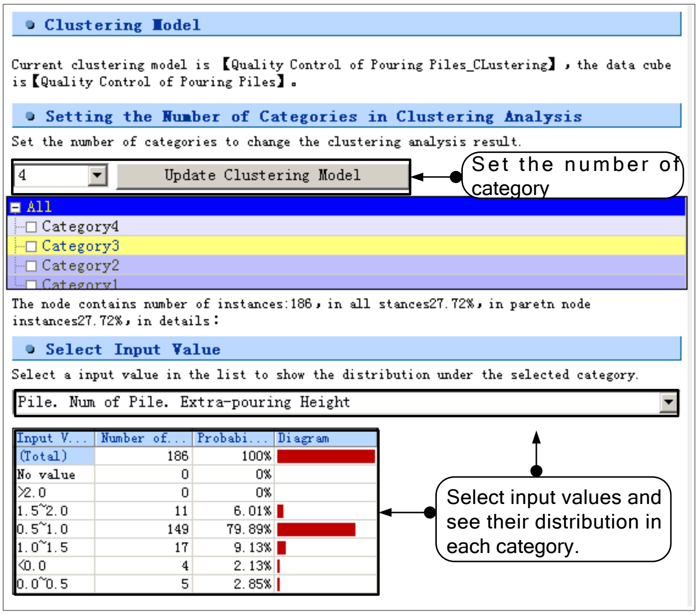

Fig. 6 Interface of clustering analysis

In this case, the number of categories was set as 2 initially and then changed, so that we can see each input value's distribution in each category to find their characteristics. Through comparison, it is found that when the number of categories was set as 4, an interesting result can be obtained as shown in Fig. 6. The corresponding characteristics of the categories are shown in Table 2. It is seen that the piles with diameter of 800 mm were divided into two categories. Among them, the extra-pouring heights corresponding to Category 3 were better controlled than that corresponding to Category 2. It means that there was a dissonance in the construction quality of the piles with diameter of 800 mm. In response to this, the management staffs may need to find out the reason and take measures to improve the situation.

Table 2 Result of clustering analysis

Category	Characteristics			
	Constructed by	Extra-pouring height	Filling coefficient	Diameter of pile
1	Urban Jinqin	1.0-1.5 m	1.00-1.05	1000 mm
2	Urban Jinqin	1.0-1.5 m	1.15-1.20	800 mm
3	Urban Jinqin	0.5-1.0 m	1.10-1.15	800 mm
4	Constructed by other contractors			

Compared with "OLAP analysis" function, "clustering analysis" function could show not only the average value of objects, but also their distribution. Thus, although it is a bit more complex to operate, it can provide further information to judge what process should be focused to control the construction quality.

4 Conclusions

The research proposes an integrated solution of decision support system for construction projects based on standardized exchanged documents. Such techniques for in-depth information analysis as data warehouse, OLAP, and data mining were utilized to develop the corresponding system Explyzer+. The system could effectively assist the management staffs to carry out the whole process of in-depth information analysis. The case study in the research showed that the functions of data mining that were provided in the system were useful for decision-making. Thus, it is feasible and effective to utilize the system to support the decision-making of management staffs in the construction projects.

References

[1] Ma Zhiliang, Li Heng, Yang Jun. Using XML to support information exchange in construction projects. *Automation in Construction*, 2004, **13**(5): 491-506.

[2] Caldas C H, Soibelman L, Gasser L. Methodology for the integration of project documents in model-based information systems. *Journal of Computing in Civil Engineering*,

2005, **19**(1): 25-33.

[3] Ma Zhiliang, Yang Jun. A study on data standard for documents in construction projects. *Journal of Harbin Institute of Technology*, 2003, **35**(Special Issue): 64-68. (in Chinese)

[4] Qin Liang. A study of the information exchange standard for construction project [Dissertation]. Beijing：Tsinghua Univiersty, 2004. (in Chinese)

[5] Ma Zhiliang, Li Hao. Study on the standardization of construction form information. *Construction Technology*, 2007, **36**(12): 36-39. (in Chinese)

[6] Zhang Jianping, Ma Tianyi. Information technology in the strategic management of construction corporations. *China Civil Engineering Journal*, 2004, **37**(12): 81-86. (in Chinese)

[7] Rujirayanyong T, Shi J J. A project-oriented data warehouse for construction. *Automation in Construction*, 2006, **15**(6): 800-807.

[8] Soibelman L, Kim H. Data preparation process for construction knowledge generation through knowledge discovery in databases. *Journal of Computing in Civil Engineering*, 2002, **16**(1): 39-48.

[9] Ma Zhiliang, Wong K D, Li Heng, et al. Utilizing exchanged documents for decision support in construction projects based on data warehousing techniques. *Automation in Construction*, 2005, **14**(3): 405-412.

[10] Gu Weihua. A system for information management and decision support in a large-scale construction project [Dissertation]. Beijing: Tsinghua Univiersty, 2005. (in Chinese)

[11] Ma Zhiliang, Gu Weihua, Li Jiulin, et al. Application of EPIMS in the construction management of the National Stadium. In: Proceedings of the First International Conference on Construction Information Technology (ICCI2004). Beijing, China, 2004: 155-165.

[12] Ma Zhiliang, He Xiaomin. Automatic generation of document templates in web-based platform systems for the collaboration of multi-party. In: Proceedings of the Ninth Conference of National Construction Information and Multimedia-assisted Projects. Weihai, China, 2006: 128-133. (in Chinese)