Life Cycle Modelling and Design Knowledge Development in Virtual Environments

Rabee M. Reffat[1] and John S. Gero[2]

[1] King Fahd University of Petroleum and Minerals, Saudi Arabia

Email: rabee.reffat@gmail.com

[2] George Mason University, USA

Email: john@johngero.com

ABSTRACT

The construction industry has been adapting information technology including virtual environments in the life cycle of buildings in terms of design, construction, operation and maintenance. The data generated using information technology within the construction industry has become increasingly overwhelming. Data mining presents an opportunity to increase significantly the rate at which the volumes of data generated through the building life cycle can be turned into useful information. This can be done by discovering patterns and correlations within a large volume of data. This Chapter presents a virtual mining environment developed to provide dynamic decision support to improve building life-cycle modelling and management.. This virtual mining environment integrates data mining with agent-based technology, database management systems, object-based CAD systems, and 3D virtual environments. A system prototype has been developed and

implemented to support the automated feed back of knowledge produced for use in building life cycle modelling, planning and decision-making.

INTRODUCTION

As the construction industry adapts to new computer technologies computerized design aids, construction, and maintenance data are all becoming increasingly available. The growth of many business, government, and scientific databases has begun to far outpace an individual's ability to interpret and digest the data. Such volumes of data clearly overwhelm the traditional methods of data analysis such as spreadsheets and ad-hoc queries. For instance, current information technology applied to facility maintenance utilises databases to keep track of information and notification of maintenance schedules. However, these databases are not well linked with interactive 3D models of buildings and are mostly presented in tabular formats. Applying techniques from a new area, called Data Mining and Knowledge Discovery, to the records of existing facilities has the potential to improve the management and maintenance of existing facilities and the design of new facilities. This will lead to more efficient and effective facilities maintenance and management through better planning based on models developed from available maintenance data, resulting in a more economical life cycle of buildings. Furthermore, designers and maintenance managers will be better equipped to achieve higher performance by utilising appropriate techniques of information technology at their workplace.

Data Mining and Knowledge Discovery in databases are tools that allow identification of valid, useful, and previously unknown patterns within existing databases (Witten and Frank, 2000; Christiansson, 1998; Frawley, 1992). These technologies combine techniques from machine learning, artificial intelligence, pattern recognition, statistics, and visualization to automatically extract concepts, interrelationships, and patterns of interest from large databases. Data mining is capable of finding patterns in data that can assist in planning. Patterns and correlations identified from data mining existing records of maintenance and other facilities management activities provide feed back and can improve future maintenance operation decision making and inform strategic planning as well as the design of new facilities. Most available computer tools for the building industry offer little more than productivity improvement in the transmission of graphical drawings and textual specifications, without addressing more fundamental changes in building life-cycle modelling and management. Virtual environments (VEs) can provide designers and facility managers with a foundation to work distributedly. Designers, building owners, facility managers and technicians can visualize and navigate the virtual building modelled in distributed virtual environments without being co-located. Information can be shared in different ways depending on the way in which and the extent to which the information must be coupled.

A virtual mining environment has been developed to integrate data mining with agent-based technology, database management systems, object-based CAD systems and 3D virtual environments. This Chapter presents a system prototype of a virtual mining environment to provide dynamic decision support for improving building life-cycle modelling and management.

LIFE CYCLE MODELLING (LCM)

The life cycle cost concept is addressed in the British Standards as 'Terotechnology' which is defined as a combination of management, financial, engineering, building and other practices applied to physical assets in pursuit of economic life-cycle costs. Life cycle cost modelling (LCM) contributes to competitiveness of the company by providing strategic planning on rehabilitation and enhanced information for decision making. LCM helps facility managers in evaluating alternative equipment and process selection based on total costs rather than the initial purchase price. The multidimensional information that LCM presents is merged from hybrid project domains such as management, engineering, as well as finance. LCM may be applied in a wide range of critical functions including: (a) evaluation and comparison of alternative design; (b) assessment of economic viability of projects and products; (c) identification of cost drivers and cost effective improvements; (d) evaluation and comparison of alternative strategies for product use, operation, test, inspection, maintenance; (e) evaluation and comparison of different approaches for replacement, rehabilitation/life extension or disposal of aging facilities; (f) optimal allocation of available funds to activities in a process for product development; (g) assessment of product assurance criteria through verification tests and their trade-offs; and (h) long-term financial planning.

With the increased use of information technology at the end of last century, the increase of information availability has become a dilemma due to inefficiencies in processing the information for decision-making. This problem becomes critical in the building industry when the high degree of complexity of work flows involved in and the accompanying uncertainty for decision making in the lifetime of a building are considered. Thus, efficiently dealing with information from different stages of a building's life cycle to improve profitability, productivity as well as strategic resource planning has become a business driver for life cycle modelling. However, the design of new buildings and facilities tends to focus on short-term cost and the immediate needs of the owner for a building that meets various business and functional requirements. Current technologies such as Computer-Aided Design (CAD) have focussed on the needs of designers to develop designs that meet design briefs that do not include life cycle design. Very little attention has been given to modelling of the life cycle costs of buildings at the design and management stages to forecast and achieve the most economical life cycle cost. There are several life cycle models available for buildings as a whole and for their component systems; although there is no one model that has been accepted as a standard, there are some areas of commonality. Life cycle cost models form predictions based on several parameters, some of which include a degree of uncertainty, such as the reliability of a part. These inputs can range from the cost of installation to the cost associated with carrying spare parts in inventory (Siewiorek and Swarz 1982). By accurately predicting failure rates and repair costs, it is also possible to compute the optimal schedule of preventative maintenance for each asset. What can be predicted and the accuracy of those predictions depends of course on the availability and accuracy of the maintenance data. Furthermore, current life cycle modelling systems fail to provide a seamless integration of hybrid information that provides users access to previously inaccessible knowledge. It has been shown in the AEC (Architecture, Engineering and Construction) industry that major factors contributing to construction quality problems include inadequate information and poor communication (Burait et al, 1992; Arditi and Gunaydin, 1998). The detection of previously undiscovered patterns in Building Maintenance System (BMS) data can be used to determine factors such as cost effectiveness and expected failure rate of assorted building materials or equipment in varying environments and circumstances. These factors are important throughout the life cycle of a building, and such information could be used in the design, construction, refurbishment, and maintenance of a building, representing a potential decrease in cost and increase in reliability. Such knowledge is significant for saving resources in construction projects.

LCM IN VIRTUAL ENVIRONMENTS

Virtual Environments (VE) are computer generated synthetic environments in which users are provided with multi-modal, highly natural forms of computer interaction. Virtual environment research is concerned with creating artificial worlds in which users have the impression of being in that world and with the ability to navigate through the world and manipulate objects in the world. Four existing life cycle modelling prototypes based on buildings with an object-oriented database in virtual environment platforms were surveyed. These projects include: "Future Home" (Murray et al, 2000); "Virtual Building Life Cycle (VBLC)" (Linnert, 1999); "LICHEE (Life Cycle House Energy Evaluation)" (CSIRO, 2003); and "Product Model and Fourth Dimension (PM4D)" (Fischer and Kam, 2000). The various benefits derived from these projects include:

- Use of virtual environments as platforms to provide an interactive interface to improve communications between all team members as well as a simulation tool in enhancing predictable strategic planning within the whole life cycle of selected buildings.
- Application of the IFC-compliant (Industry Foundation Class) object oriented database in standardizing the data exchange and facilitating the manipulation and reusability of project information.

- Linkage of maintenance data to 3D CAD model provides the potential of future development of intelligent life cycle analysis and control capability.

However these prototypes do not provide the capability of data mining of the hybrid data gathered from different stage of a building's life cycle, and do not provide performance-gaining life cycle analysis.

A DATA MINING APPROACH TO LCM

Data mining has been defined as the "nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (Frawley et al, 1992). Mining data enables people to understand how systems that were once thought to be completely chaotic have predictable patterns (Peitgen et al, 1992). Through data mining, patterns and causal relationships behind apparently random data in AEC projects can be found. By applying data mining to identify novel patterns, project managers will be able to build knowledge models that may be used for recurrent activities of on-going construction projects, as well as for a future project activities, and avoid unanticipated consequences (Soibelman and Kim, 2002). Data mining also presents the potential to address the problem of transforming knowledge implicit in data into explicit knowledge for decision making. In contrast to traditional methods of statistical data analysis, data mining is an automated process that discovers trends and patterns without the need for human intervention. Data mining takes input variables whose relevance may not be obvious to a designer but which becomes evident as a result of this process. In addition, data mining makes no prior assumption about the probability distribution of the input variables (Gaussian, Poisson, etc), as is required in statistics, and is therefore more robust and general. However, like other methods, the process of transforming the data into a format suitable for knowledge discovery is not automated and has a large impact on the results obtained. Thus, the approach is based on a comprehensive view of the building management problem. It views the process of building design, maintenance, and replacement as a process generating an enormous amount of information. While current practice addresses parts of this information generation and management, our approach attempts to account for the life cycle flow of this information.

The costs of designing and building structures are much smaller than the costs of operating a building or other structure over the course of its life span. The knowledge that becomes available through data mining enables a building owner to make important decisions about life cycle costs in advance, thereby significantly affecting and improving design decisions. The rich set of building data that is created during the design and documentation phase of the building remains relevant even after the building is constructed, and the data only becomes richer as maintenance data is added. Architects, interiors designers and engineers, as well as contractors, marketing and sales personnel, building managers and owners can extract information from the databases for the building's renovation, maintenance, and operation. Figure 1 outlines the proposed model of the flow of information in building design and maintenance. The bold arrows depict the functionality provided by our proposed approach while the dashed arrows describe the scope of present approaches to building information management.

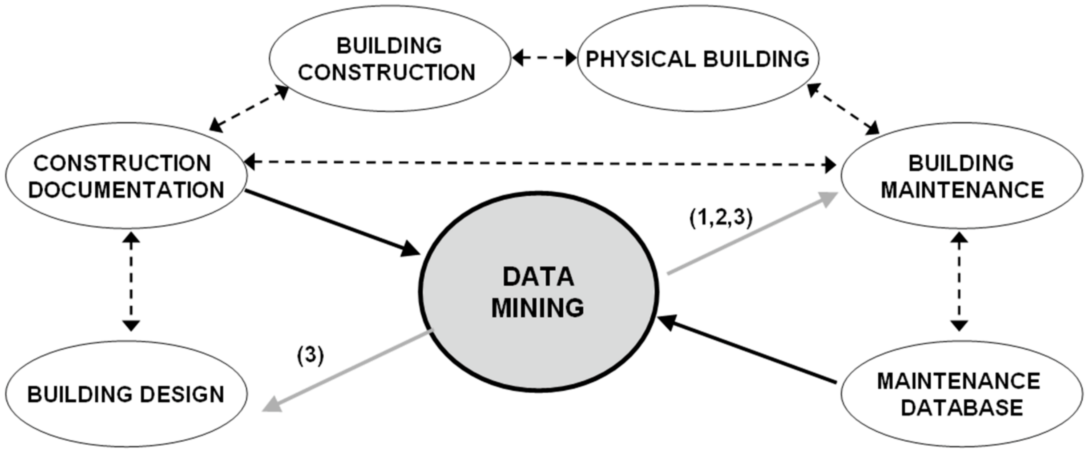


Figure 1. Integrating data mining within the life cycle of building information management.

Data mining techniques can be used effectively on data stored in a building maintenance system by creating knowledge that can be used in future management and design decision making. Knowledge that implicitly resides in building maintenance system databases includes information about:

1. components that frequently need maintenance and therefore need to be inspected carefully,
2. historical consequences of maintenance decisions that may inform future decisions, and
3. components of buildings that significantly determine maintenance cost and therefore may inform future building designs, as well as refurbishment of the building in question.

This information can be extracted using data mining techniques and used to improve all phases of the building life cycle, both for current and future buildings, as indicated by the numbers (1, 2, 3) and (3) marked on the arrows in Figure 1, which correspond to those listed above. On the other hand, the processing and transformation of data into a format suitable for data mining is an important step in a successful application of data mining and the stages for extracting knowledge from data using Data Mining techniques are shown in Figure 2

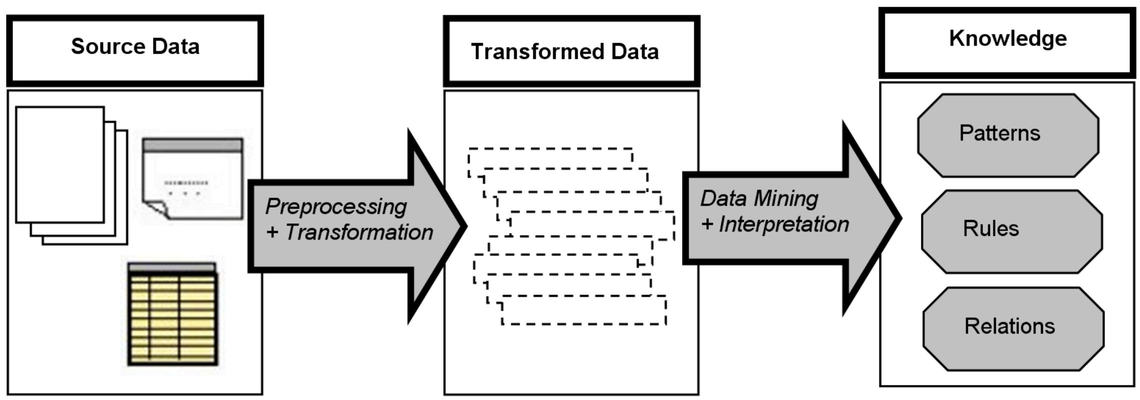


Figure 2: Stages for extracting knowledge from data using Data Mining techniques.

The data mining approach for life cycle modelling in this work views the process of building design, maintenance, and replacement as a process that generates high volumes of information. While current practices address only parts of this information generation and management, this approach attempts to account for the life cycle flow of this information. The integration of data mining within the process of information flow provides the opportunity to increase the value of building data and to feedback and improve the processes of building design and maintenance.

Incorporating Data Mining in Virtual Environments for LCM

A significant need exists for the application of new techniques and tools to automatically assist humans in analysing the increasing volume of data for useful knowledge. The increasing use of databases to store information about facilities, their use, and their maintenance provides the background and platform for the use of data mining techniques for future projections. The current technology for facility maintenance uses databases to keep track of information and for notification of maintenance schedules. These databases are not well linked with an interactive 3D model of the building and are generally presented in tabular form.

A VIRTUAL MINING ENVIRONMENT FOR PROVIDING DYNAMIC DECISION SUPPORT

The virtual mining environment promotes real-time support and multiple team participation and involvement. The virtual mining environment can be remotely accessed synchronously by different users who will be aware of the presence of others and communicate with them. Users might mine the same or different building element based on their focus and interest. Each user might be looking at different building assets and using different mining and discovery techniques than others within the same virtual environment. Within the virtual mining environment data-mining techniques are utilized to discover rules and patterns of useful knowledge from the maintenance records of a building to help improve the maintenance management of existing and future buildings. Although it may sound at first appealing to have an autonomous data-mining system, in practice, such a system would uncover an overwhelmingly large set of patterns, and most of the patterns discovered in the analysis would be irrelevant to the user. Therefore, it is important to provide a user focused approach to mine the maintenance data of buildings.

The virtual mining environment prototype system described in this chapter includes: an object oriented 3D CAD model of a building modelled in the ArchiCAD package, and a maintenance database in a standard SQL (standard query language) format. The architecture of this virtual mining environment has been developed to include three agents: interface, maintenance and filter agents, as illustrated in Figure 3. The roles of these agents include:

(a) The appropriate mapping between the building assets of the building model in the virtual environment is maintained by the maintenance agent that connects data contained within the maintenance database with data contained within the Express Data Manager (EDM) database via the virtual environment, Active Worlds.

(b) Linking data mining techniques to building models in a virtual environment (Active Worlds) is achieved via the maintenance agent that accesses the maintenance database and applies its mining algorithms on it.

(c) Linking knowledge development with the building model in virtual environments is carried out by the filter agent that assists in improving maintenance management by providing life cycle implications as feedback whenever building assets (mechanical and electrical elements) are selected in the building model in the virtual environment.

Feedback of useful knowledge can be discovered by the maintenance agent in the application of the four data mining techniques and algorithms (aimple k-means, apriori, ID3, and C4.5) (Witten and Frank, 2000) in order to discover various classifications of knowledge. The data mining algorithms and the link between its knowledge development and the building model in a 3D virtual environment has been fully implemented. The use of the virtual mining environment requires the utilization of the following four phases progressively:

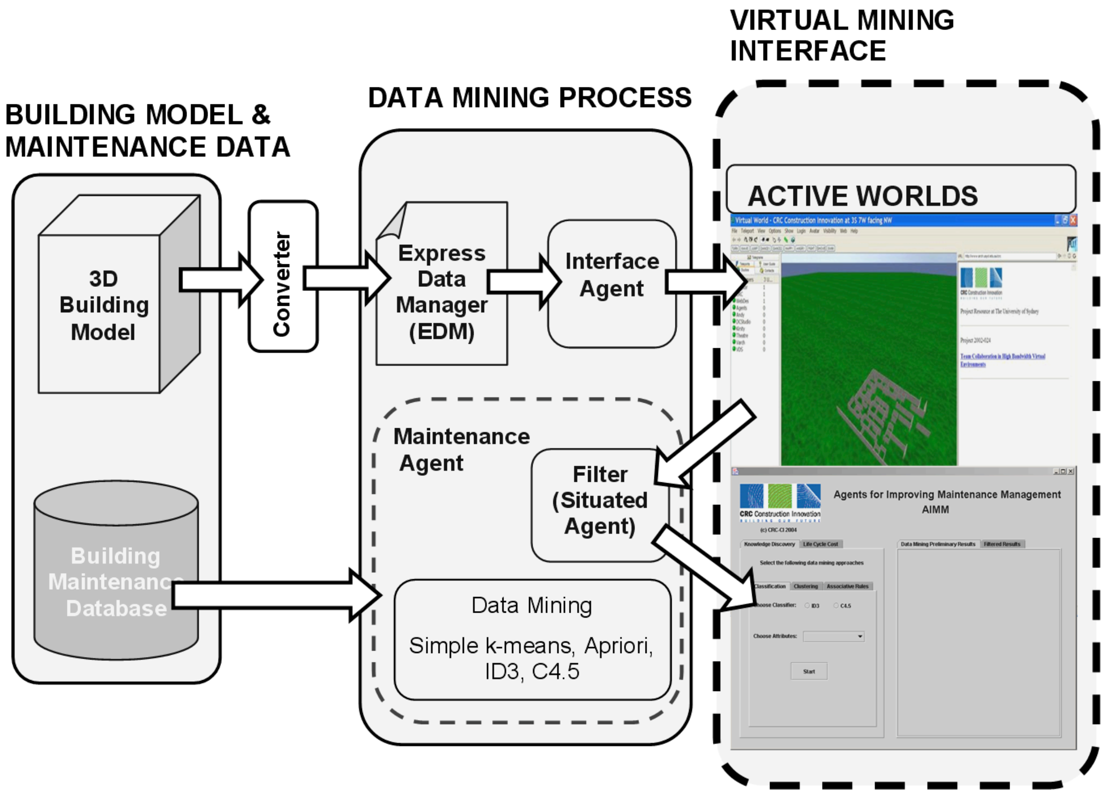


Figure 3. Architecture of the virtual mining environment.

- Phase 1: involves the manual pre-processing of data, which removes noisy, erroneous and incomplete data to derive important attributes from original raw data. For example, the raw text description of time of work orders "1/12/2001" which need to be converted to a meaningful attribute such as "month". Moreover, various "testing" algorithms are run through the maintenance data to find out the suitable data mining approaches. As a result of Phase 1, the quality of the data is improved.
- Phase 2: adopts the EDM interface agent developed by Maher and Gero (2002) in converting IFCs (Industry Foundation Classes) object model into a Renderware (RWX) format, so as to be available in the virtual environment. The virtual

environment provides a collaborative multi-user interface and more importantly, a means for the user to walkthrough 3D object model at a real time. The user is able to navigate and select a building asset type to explore useful knowledge.

- Phase 3: instantiates the maintenance interface agent and the maintenance agent. Once the user decides to select a certain building asset the maintenance interface agent is invoked to load related data from database. The maintenance agent performs data mining on the selected asset type. The four mining algorithms that have been implemented in the system prototype are: clustering using "simpleKmeans", associative rules learning in "apriori", classification using "C4.5" and "ID3".
- Phase 4: A filter agent is activated. This is a software agent that performs post-processing of the mined results. The filter agent filters out irrelevant patterns based on the heuristic rules.

The Virtual Mining Environment Interface and User Scenario

The user is able to navigate a 3D model within a real time virtual environment as shown in the top left of Figure 4. The user is able to instantiate the virtual mining environment prototype system by clicking on the desired building asset or component, invoking the main Maintenance Interface as shown in the bottom right of Figure 4.

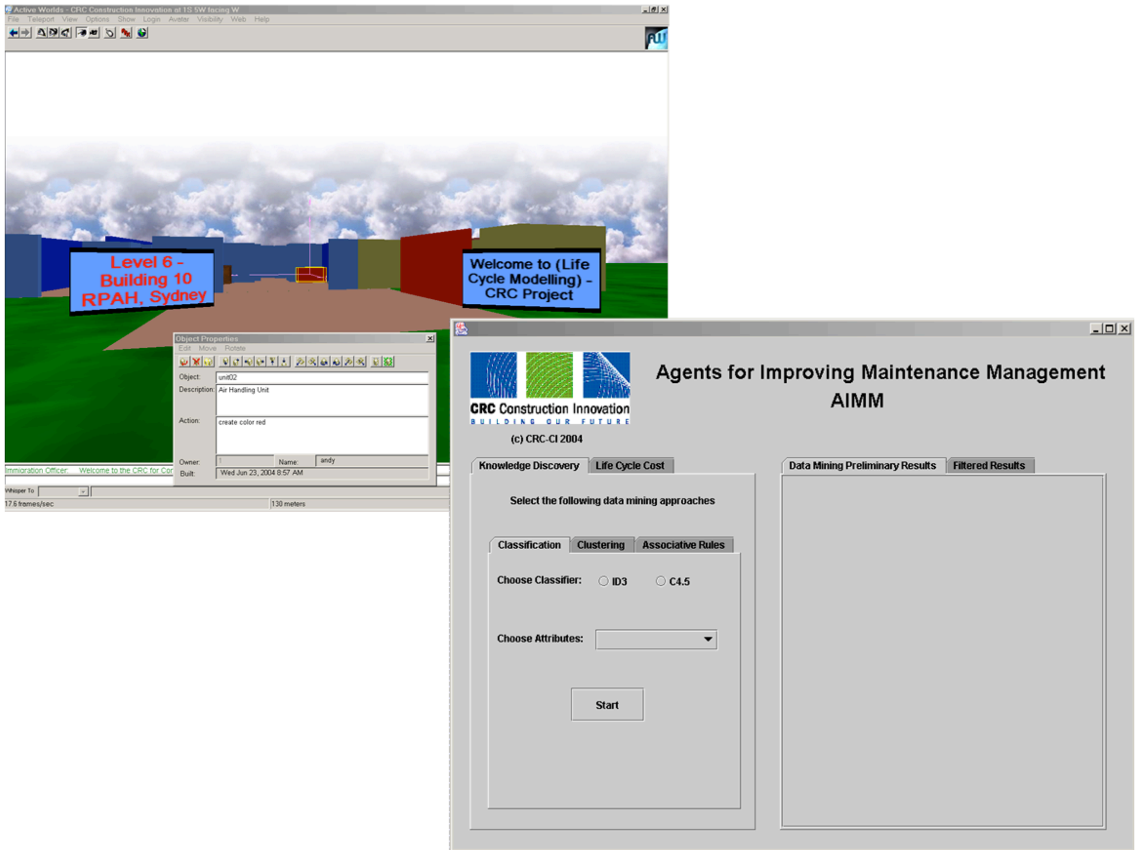
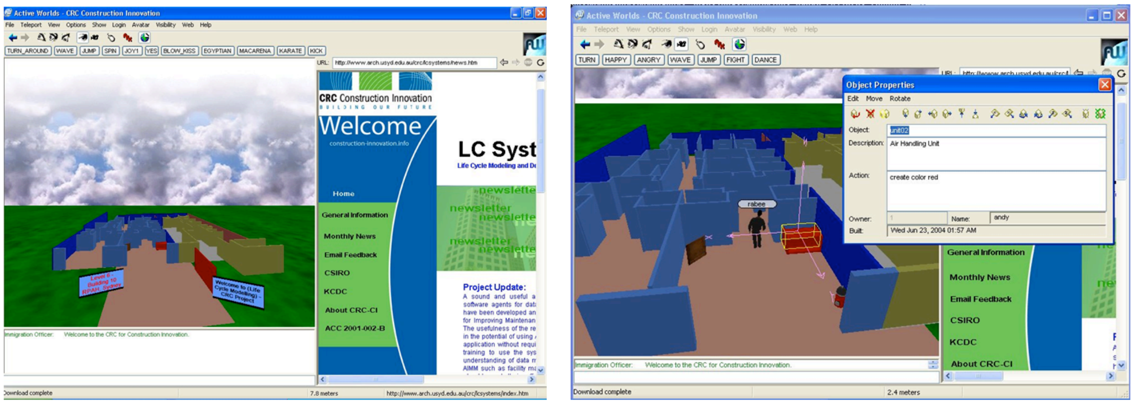


Figure 4. Selecting an asset type in Active Worlds instantiates the maintenance interface agent.

Once the interface agent is activated it pops up the knowledge discovery panel which has three stacked sub-panels: (i) classification, (ii) clustering and (iii) associative rules. These panels provide a range of ways for using each different algorithm. This provides the user with greater flexibility and scope since the user may test a variety of data mining approaches for each type of algorithm. On the right hand side are another two stacked panels that are dedicated to reporting results. Results are reported in two ways. The panel named :Data Mining Preliminary Results" displays the results of the chosen algorithm in their "raw" form. The panel named "Filtered Results" displays the results in their interpreted form using domain derived heuristics. The overall data mining interface is illustrated in Figure 5(a) and illustrated the hierarchy of stacked panels for the different data mining scenarios.

In this scenario, an Air Handling Unit (AHU) is selected as the building component that a user wishes to apply data mining to. The following sequence of actions is then followed:

- The user navigates the building in a real-time and online 3D virtual environment as shown in Figure 5(a);
- Once the user selects a building asset type such as the Air Handling Unit the object property window pops out describing general information of the selected object as shown in Figure 5(b).;
- Then, the interface agent is invoked and the main window pops up to allow selection of algorithms as illustrated in Figure 6(a);
- User explores a variety of data mining algorithms and chooses the desired algorithm and runs it as shown in Figure 6(b);
- The maintenance agent running the algorithm is invoked and results are reported first in the Data Mining Preliminary Results panel as illustrated in Figure 6(c);
- User selects Filtered Results in order to access post-processed knowledge and an example of filtered knowledge is shown in Figure 6(d).



(a) (b)

Figure 5. (a) The primary interface of software agents prototype system [AIMM] in an interactive network multi-user environment; and (b) The user selects a building asset type (the Air Handling Unit) and an object property window pops out describing general information of the selected object.

Data mining techniques assisted in identifying critical cost issues. For instance, discovering that corrective maintenance accounts for approximately 55% of all work orders implies a high level of unplanned maintenance that contributes to increasing the operational cost. The maintenance services required for the air conditioning system were related to thermal sensation complaints (too_hot 32%, too_cold 28%, not working 7.5%; total 67.5%). Hence, applying data mining techniques assists facility and building managers to identify the crucial maintenance issues and directs the improvement of strategic planning to add value to the life cycle of buildings.

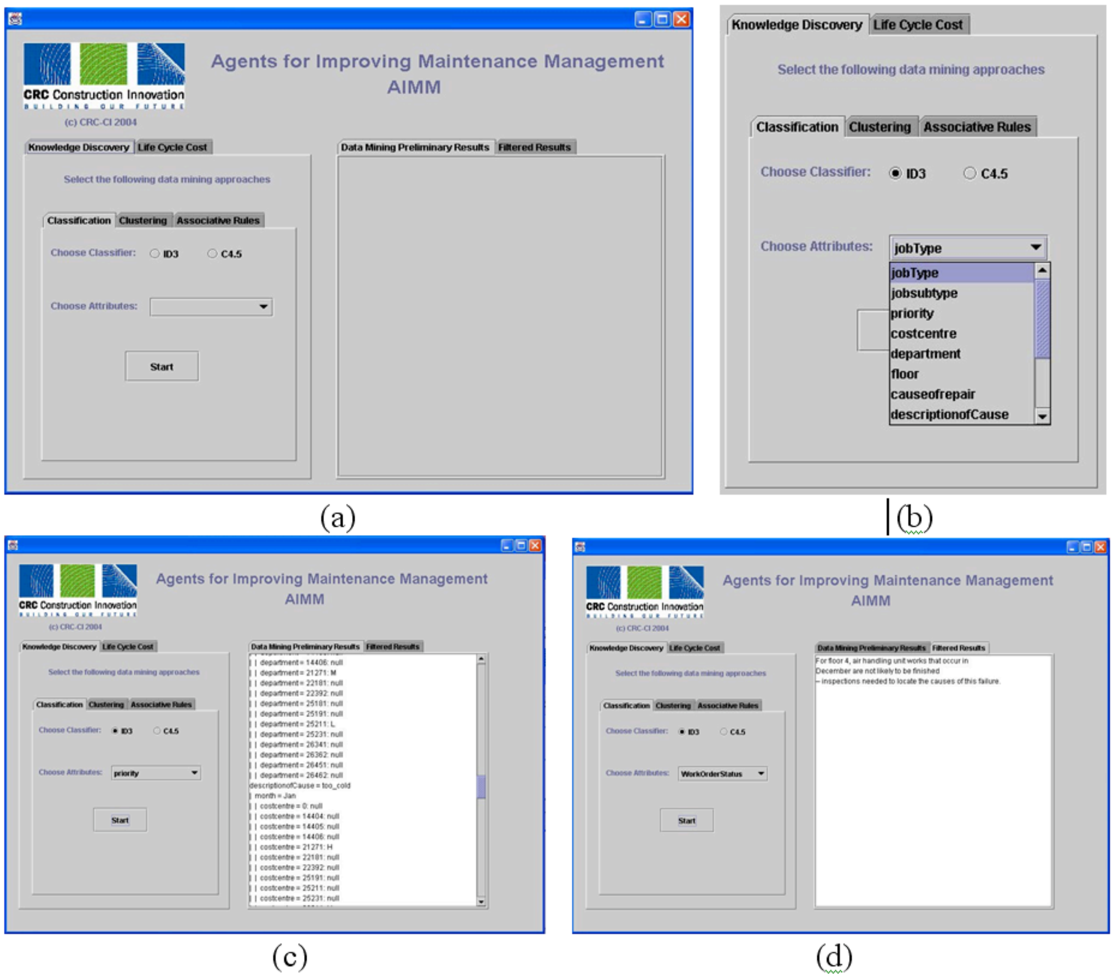


Figure 6. (a) The AIMM prototype system is instantiated once a building asset type has been selected; (b) Data Mining techniques and different attributes for the user to choose from based on focus and interest; (c) Preliminary results of applying the ID3 with the "Priority" attribute on the maintenance data of Air Handling Unit; and (d) An example of the filtered knowledge presented to the user from the preliminary results of applying the ID3 with the "Work Order Status" attribute on the maintenance data of Air Handling Unit.

Other benefits include constructing predictive plans based on correlations obtained from applying data mining techniques on the maintenance data sets of buildings. For instance, considering the role of potential correlations between seasons and malfunction rates in guiding the allocation of maintenance resources. Also, investigating any abnormal phenomenon discovered from the maintenance data set such as “all outstanding works took place in December”. An investigation is required to study the relationship between the cause of increasing the outstanding maintenance jobs taking place in December the Christmas holiday or any the other causes. Appropriately addressing this problem will lead to better activities to improve the maintenance management of existing facilities and will guide the design of future facilities (Reffat et al, 2004a and 2004b; Reffat and Gero, 2005).

DISCUSSION

The development of data mining agents of facilities and building maintenance data in a 3D virtual environment provides a new approach for improving the maintenance and management of building facilities and guiding future design decisions. Virtual environments of building models offer the opportunity for the user to navigate through the model, to manipulate and to interact with its objects. The integration of facilities databases with interactive 3D virtual environments containing building models and data mining techniques provides a visual modelling tool for the simulation and projection of the financial and physical impact of maintenance, refurbishment and major replacement and extension of a building and its components over its live cycle. The virtual mining environment has the potential to change the conventional approaches of applying data mining, and to provide a cutting edge technology to transform the way decision support for building maintenance is carried out presently and in the future. The virtual mining environment facilitates real-time assistance for decision making and provides an interactive platform for designers, building owners, facility managers and technicians to communicate and interact collaboratively and virtually within a 3D real time and multi user virtual environment. The building model presented in the real-time multi-user virtual environment is composed of sets of 3D building elements. Building elements include walls, doors, floors, windows, roof, and mechanical and electrical equipments. Each building element is linked to both a knowledge-base (that includes the maintenance records of that element), and to a data-mining agent triggered based on user's request to mine the knowledge-base and provide useful knowledge that help enhancing the decision making process to maintain and mange building facilities.

Data mining of building maintenance can help to discover: procedures that reduce future failures; repairs or maintenance operations that are being executed improperly; ways to improve repairs that reduces subsequent down time; undocumented methods being used by experienced personnel that result in reduced down time; advance notice of likely failures before failures occur. Such discoveries can be used to modify building maintenance and repair procedures thereby reducing downtime, increasing uptime, and significantly reducing the costs of maintenance and repair. Furthermore, the knowledge-base of building elements is not static since all daily maintenance actions that took place are updated in the knowledge-base. Hence, the knowledge acquired by data-mining agents is active and dynamic. Therefore, the virtual mining environment facilitates providing a live, active and dynamic decision support on building maintenance to improve its maintenance management and the building life-cycle.

ACKNOWLEDGEMENTS

The research presented in this paper is funded by CRC-Construction Innovation, Project No.: 2001-002-B "Life Cycle Modelling and Design Knowledge Development in Virtual Environment". The industrial maintenance data is provided by Central Sydney Area Health Service, Royal Prince Alfred Hospital, NSW, Australia. The assistance of Lan Ding, Julie Jupp, PakSan Liew and Wei Peng in the development of this project is gratefully acknowledged.

REFERENCES

Arditi, D. and Gunaydin, M. (1998) 'Factors that affect process quality in the life cycle of building projects', *Journal of Construction Engineering and Management, ASCE*, 1998, 124 (3): 194-203.

Burait, J., Farrington, J. and Bedbetter, W. (1992) 'Causes of quality deviations in design and construction', *Journal of Construction Engineering and Management*, 118(1): 34-49.

Christiansson, P. (1998) 'Using knowledge nodes for knowledge discovery and data mining', in I. Smith, (ed) *Information Technology for Design, Collaboration, Maintenance, and Monitoring*, Berlin: Springer-Verlag, 48-59.

CSIRO (2003) 'Annual Report 2002-2003 Commonwealth Scientific and Industrial Research Organisation, Online Available HTTP: <http://www.csiro.au/files/files/p2kv.pdf> (accessed 20 February 2008).

Fischer, M. and Kam, C. (2002) 'PM4D Final Report, CIFE (Center for Integrated Facility Engineering)', *Technical Report Number 143*, Stanford: Stanford University,

Frawley, W., Piatetsky-Shapiro, G. and Matheus, C. (1992) 'Knowledge discovery in databases: An overview', *AI Magazine*, 13: 57-70.

Linnert, C. (1999) 'Virtual building lifecycle: visualization of lifecycle data on a virtual model of a building in a 3D-CAD environment', *Diploma Thesis*, Karlsruhe: Institute for Industrial building production, University of Karlsruhe.

Maher, M. L. and Gero, J. S. (2002) 'Agent models of virtual worlds', *ACADIA 2002: Thresholds,* California State Polytechnic University, Pomona, CA, 127-138.

Murray, N., Fernando, T., Aouad, G. (2000) 'A virtual environment for building construction', *17th ISARC*, 1137-1142.

Peitgen, H. O., Jurgens, H. and Saupe, D. (1992) *Chaos and Fractals: New Frontiers of Science*, New York: Springer-Verlag.

Reffat, R. , and Gero, J. (2005) 'A virtual mining environment for providing dynamic decision support for building maintenance', *Proceedings of the 23rd Conference on Education of Computer Aided Architectural Design in Europe (eCAADe2005): Digital Design-The Quest for New Paradigms*, Lisbon, 589-596.

Reffat, R., Gero, J. and Peng, W. (2004a), 'Improving the management of building life cycle: A data mining approach', *CRC Research Conference*, Brisbane.

Reffat, R., Gero, J. and Peng, W. (2004b) 'Using data mining on building maintenance during the building life cycle', Proceedings *of the 38th Australian & New Zealand*

Architectural Science Association (ANZASCA) Conference, Tasmania: School of Architecture, University of Tasmania, 91-97.

Siewiorek, D. P., Swarz, Robert S. (1982) *The Theory and Practice of Reliable System Design*, Bedford: Digital Press.

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

Witten, I. and Frank, E. (2000) *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, San Diego: Morgan Kaufman.