Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 111 (2017) 417 – 425

Energy Procedia

8th International Conference on Sustainability in Energy and Buildings, SEB-16, 11-13 September 2016, Turin, ITALY

Case-based reasoning and text mining for green building decision making

Xue Xiao*, Martin Skitmore, Xin Hu

Queensland University of Technology, 2 George St, Brisbane, 40058, Australia

Abstract

There are great benefits to be obtained by sharing previous experiences in meeting the needs of the standard evaluation systems for green building around the world. To date, there are no existing methods available that enable this to take place in a systematic way. This paper addresses the issue by developing a green building experience-mining (GBEM) model that enables previous green building solutions to be adapted for a new situation. A database of 10 cases is used to demonstrate and evaluate the effectiveness of the GBEM model. The results confirm the model's potential to facilitate users in the selection of the solutions when addressing new green building challenges.

Peer-review under responsibility of KES International.

Keywords: case-based reasoning; text mining; green building; decision making;

1. Introduction

Green building is attracting increasing attention for its role in environmental preservation as well as the economic benefits involved. The popular concern for green building reflects the potential importance of construction for environmental conservation [1]. Small improvements in the sustainability of buildings, or in the energy efficiency of their design and construction, have a significant influence on the amount of their energy consumption. Therefore, the implementation of green building is significant because it has become an increasing important attribute of economic activities concerning methods of production [1].The development of green buildings has a large effect on city development, as evidenced by several studies of the relationship between green building and urban development [2-4].

* Corresponding author. Tel.:+61-0449-584-746;
E-mail address: xue.xiao@hdr.qut.edu.au

Peer-review under responsibility of KES International.
doi:10.1016/j.egypro.2017.03.203

Given the significance of green building, the promotion of sustainability principles for building has resulted in numerous associated practices. Most significant are the various rating systems that now exist in many parts of the world today, the main ones of which are the UK Building Research Establishment's (BRE) Environmental Assessment Methodology (BREEAM) established in 1990 and now used in more than 50 countries around the world; the U.S. Green Building Council's (USGBC) Leadership in Energy and Environmental Design (LEED) Green Building Rating System, instigated in 1994 as the nationally accepted benchmark for the design, construction and operation of high performance green buildings; the Australia Sustainable Energy Development Authority's (SEDA) Australian Building Greenhouse Rating Scheme (ABGRS) launched in 1998 as the first Australian comprehensive green building assessment system, which evaluates energy consumption and carbon dioxide emissions; and the Ministry of Housing and Urban-Rural Development of China's 2005 national "Green Building Evaluation Standard", which covers all types of civil buildings from previous residential buildings and public from the design phase thru to building operation. The German sustainable building council (GeSBC) was formed in 2007 and developed a certification system tool for new office buildings, which covers more than fifty criteria with quantification rules and target values for new construction office buildings. The French eco-building rating system known as High-Quality Environment (HQE) was established by the non-profit Association HQE in 1992. The first Italian association of architects promoting green building awareness, the Associazione Nazionale Architettura Bioecologica (ANAB), was founded in 1989. In March 2008, the building energy certificates have been issued throughout the country for new buildings and for property transfer of existing medium and large buildings in Italy. Besides the Europe, several hot climate countries have made effort to attempt to follow the international green movement. For example, the Green Building Standard SI 281 (Israel) was founded in 2005 and issued the Israeli Green Building Standard IS 5281, a point scoring system. Pearl Building Rating System (PBRS) (UAE) is developed by The Abu Dhabi Urban Planning Council (UPC) for new construction buildings in 2007.Qatar Sustainability Assessment System (QSAS) is a green building certification system developed for the State of Qatar by the T.C. Chan Centre for Building Simulation and Energy Studies at the University of Pennsylvania.

One outcome of the application of these standard criteria ranking or evaluation systems is a large amount of experience with their use. This is a common situation in the construction industry, where experience-oriented practitioners are accustomed to addressing issues and problems by utilizing their accumulated experience and knowledge [5]. Previous experience is therefore highly important because it provides a valuable reference for current construction problems [6]. Applying this process to green building in an organized way involves establishing a database of associated experiences and developing a means of systematically reusing the information contained therein. The national standard evaluation system provides a model of the structure of the database. The system's requirements for each ranking level can be used to divide the experiences of building green construction into similar levels. Therefore, the aim of this paper is to build a green building experience-mining (GBEM) model with case-based reasoning and text mining. Case-based reasoning (CBR) is a method providing decision-makers with a framework to solve current problems by recalling and reusing previous knowledge and experience [7]. The term 'case' refers to a previous concrete situation, including problems, solutions and outcomes [7]. In order to turn the experience in the format of text to be systematically descriptive, this paper utilizes the text mining techniques that enable to translate the semantic questions involved into theoretical problems. Text mining is an emerging technology that attempts to extract meaningful information from unstructured textual data[8]. It can be seen as an extension of data mining to textual data. The reminder of this study is comprised as follows: first, in the literature review section, the background of CBR and green building are introduced respectively as well as the methodologies used for the model are explained. Second, the GBEM model is proposed using the aforementioned methodology. Thirdly, a case study is conducted to validate the model's effectiveness. Lastly, the findings and contribution of this research are discussed in the contribution section.

2. Literature Review

2.1. Case-based reasoning

CBR is a problem-solving process for addressing current issues by recalling and reusing previous knowledge and experience [6]. As a branch of the artificial intelligence, CBR is extensively utilized in the experience- or knowledge-mining domain because of its capability of addressing problems intelligently through mimicking the human mind [5]. Despite the unique characteristics of every building, the methods utilized in their construction are similar[9]. This suggests that successful green building practices adopted in previous projects can be equally applied in new projects, providing important opportunities for the application of CBR [10] . The utility of construction management (CM) CBR in experience mining is widely acknowledged and is attracting increasing interest in applications such as decision-making, cost estimation and procurement selection [11-13].

The CBR process involves two stages: problem formulation and the CBR circle, as shown in Fig.1. Problem formulation consists of three steps comprising case representation, case indexing and case storage. Case representation is the conversion of the problem into a descriptive format for the CBR circle; case indexing involves assigning characterizing attributes to the cases; and case storage is the process of establishing the structure of the case base in preparation for the next step of the CBR cycle. The CBR cycle consists of four processes of retrieve, reuse, revise and retain [14, 15]. Retrieve is the process of searching for previous cases that best match the new problem from the case base; reuse is the process of grasping and utilizing the experience from the retrieved case to the new problem; revise is the process of adjusting the solution of the retrieved case to match the situation of the new case; and retain is the process of updating of the case base by storing the revised solution of the new problem [15].

The CBR process involves two stages: problem formulation and the CBR circle, as shown in Fig.1. Problem formulation consists of three steps comprising case representation, case indexing and case storage. Case representation is the conversion of the problem into a descriptive format for the CBR circle; case indexing involves assigning characterizing attributes to the cases; and case storage is the process of establishing the structure of the case base in preparation for the next step of the CBR cycle. The CBR cycle consists of four processes of retrieve, reuse, revise and retain [14, 15]. Retrieve is the process of searching for previous cases that best match the new problem from the case base; reuse is the process of grasping and utilizing the experience from the retrieved case to the new problem; revise is the process of adjusting the solution of the retrieved case to match the situation of the new case; and retain is the process of updating of the case base by storing the revised solution of the new problem [15].

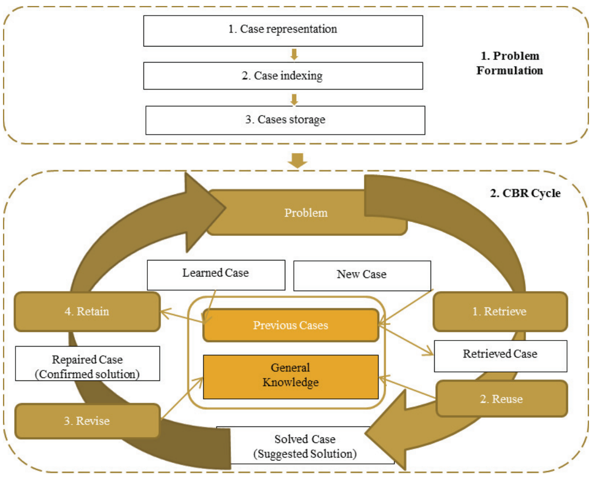

Fig.1. Process of CBR (adapted from Aamodt and Plaza (1994) and (Kim, 2011))

As Fig.1 illustrates, CBR techniques have three characteristics: the solution for a new building can be obtained from existing experience [6]; the suitability of the match between the retrieved solution and the current case is heavily influenced by its method of retrieval - the indexing strategy [12]; and the size of the case base is a significant factor affecting the efficiency of CBR-based models [16] .

2.2. Brief overview of green building

Green building, also known as sustainable building and sustainable construction, most comprehensively addresses the ecological, social and economic issues of a building in the context of its community. The term 'green building' refers to the quality and features of the actual structure created using the principles and methodologies of sustainable construction [17]. It is defined in the American Society of Testing and Materials (ASTM) Standard E2114-06a as "a building that provides the specified building performance requirements while minimizing disturbance to and improving the functioning of local, regional and global ecosystems both during and after its construction and specified service life" [18]. This definition is closely related to the definition of sustainable construction. In 1994, the International Council for Building (CIB), an international construction research networking institution, articulated the definition of sustainable construction as "creating and operating a healthy built environment based on resource efficiency and ecological design." There are seven basic principles of sustainable construction that are also the basic ecological principles for green building comprising : (1) reduce resource consumption; (2) reuse resources; (3) use recyclable resources; (4) protect nature; (5) eliminate toxics; (6) apply life-cycle costing; and (7) focus on quality.

The promotion of sustainable principles for building has resulted in numerous related studies. Hu et al (2010), for example, study the practice of the sustainable design of green building in Chinese campuses, using the Logistical Engineering University as an example for demonstration the combination of design and ecological impacts [19]; Stankovic et.al (2014) compare the lightning design criteria of green building certification systems as guidelines for implementing of green building in Serbia [20]. Azizi et. al analyze the management practices used in Malaysia to achieve energy efficiency performance, including the planning, procurement and monitoring phases [21]. From those documents, successful cases have been identified from the perspective of putting sustainable principles into practice.

The above discussion illustrates the popularity of experience mining for sustainable construction. For the systematic treatment of accumulated experiences relating to green building, it is necessary to retain the associated information in an organized way. Establishing the database is an urgent primary task because it enable previous good practice to inform current green building problem solving, even though the sustainability requirements of the previous projects may not be exactly the same the new one. In other words, information concerning previous cases needs to be available in such a way as to provide as many potential alternative solutions as possible as well as offering an accurate and targeted experience base as possible.

3. The GBEM model

As Fig.1 illustrates, the development of a framework of experience mining involves two main steps: firstly, problem formulation, i.e., the description of the problem and the establishment the database for experience mining; secondly, the CBR cycle or process of identifying a suitable case in the database. This involves the four steps of case representation, case indexing, case storage and the CBR circle.

3.1. Case representation

Case representation involves the identification and representation of a practice and is a process to decide the amount of the information each case should contain and the way in which it is represented. It is important to organize existing cases in a way that contains basic components useful for solving new problems. There are three main elements involved in a typical practice case: (1) description of the problem; (2) the solution adopted; and (3) the outcome of the solution [22]. In this paper case representation refers to the process of the identification of three aforementioned elements. The research problem is how to reuse the existing experience as an aid to solving green building problems. 'The solution adopted' is the good practice the previous case adopted. 'The outcome of the solution' is the result of the good practice such as the energy efficiency of the green building.

3.2. Case indexing

Case indexing is the process of translating practical content into an organized form by using a suitable index. According to Chua et.al, the effectiveness of the model is significantly influenced by the indexing strategy [23]. Because experience mining usually operates on a general domain platform [22], the database structure needs to follow the principles of that model. There are several types of the general domain model. CHEF, for example, is used in the planning domain that creates new plans by recalling old plans that worked well under similar circumstances and modifying them to fit the new situation [24]; while HYPO works in the law domain in a similar manner [25]. Both of are based on the CBR principle. In our case, the CBR model is adopted to establish the framework for experience-mining green building, with text mining used to calculate the similarity between the contexts. The process of the case indexing enables the records of existing case practice to provide useful information in an organized way to help the user better understands the problem.

Conventional attribute selection methods include literature review, questionnaire and expert subjective judgement. Literature review helps to summarise the previous information but has limitation in the suitability. Questionnaire and expert opinion offer subjective feedback but the later provides the more in depth information. Therefore, expert subjective judgement is adopted as the case indexing method in the current paper. The availability of the information is also taken into consideration given the privacy requirement of the company. After going through the context document and deep interview with the expert, the preliminary case index are decided, including the project information, result of practice, practice of green building and others.

3.3. Case storage

Case storage is final step in the establishment of the case base. According to the aforementioned case indexing, the first level components are resolved into more detailed components. Fig. 2 illustrates the structure of the components and their relationship, comprising two main types of indices: (1) the input index, including project information and results of practices; (2) the output index, including the project experience, with the practices belonging to the same category grouped hierarchically. For example, if user intent to search for some experience about the green condition of an office building, he should use green condition and office building as the input. The result will show how the previous experience on building concept, design, construction and equipment system will help.

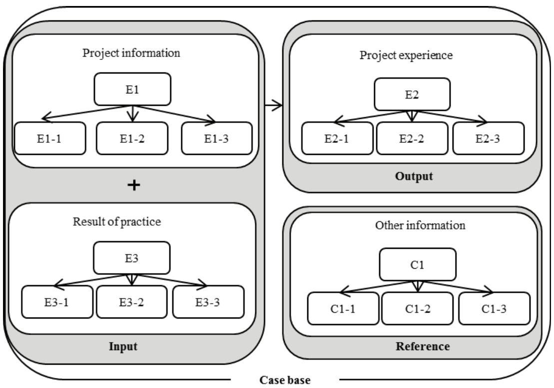

Fig.2. Case storage structure of the GBEM model

3.4. CBR cycle

The CBR circle starts with inputting a new problem and ends with the identification of the most appropriate solution for the given problem. The input information includes the project background and goal, such as the expected green rating or the energy conservation. The model then searches for the best-matched case to the new case. There are several ways of doing this, including (1) categories, (2) semantic relations and (3) index pointers [26]. All of these three methods are used in the GBEM to locate the most appropriate practice. After calculating the similarity between the retrieved case and the current case, the best case is selected, so that the most appropriate experiences of the previous case are identified to solve the current problem.

The GBEM model offers a way to search for the best practice within a massive database and is designed to address the challenge of how this can be used to realize the expected goal of a new project. Text-mining techniques are combined with the CBR model to address the problems encountered in dealing with semantic relations during the process of problem formulation and the CBR circle. The application of text mining to the CBR model is a dynamic cycle and requires continuous evaluation and examination to refine the results [27]. This involves several steps: (1) pre-processing, to narrow the case base to the specific field of enquiry – green building in this context; (2) applying text mining, involving the use of different text mining algorithms including concept extraction, categorization and clustering; and (3) evaluating the text mining result and taking action. After the automatically generated result of text mining, the consequence is carefully examined for relevance and appropriate adjustments made to enhance the accuracy and effectiveness of the model reusing the previous practice for the current problem.

4. Case study

A case study is used here to demonstrate the use of the model. An extensive search for cases of green building practice was conducted in order to obtain sufficient documents for the database form the China Academy of Building Research (CABR), which is the nation's largest and most comprehensive institution in terms of advanced construction technology and sustainable application in China. These cases are considered to provide a sufficient number for the purposes of demonstration and testing. The pre-processing phase consisted of going through a context document of more than 50 building cases in China, with 10 eventually being identified as related to the "green building experience".

Table 1. Template for case indexing green building practices

Level 1	Level 2	Level 3
1. Project information	1.1 Project name	/
	1.2 Project type	/
	1.3 Green building rating	/
	1.4 Project size	/
	1.5 Region	/
2. Result of practice	2.1 The use of renewable energy	2.1.1 The power
		2.1.2 The water
	2.2 The greening condition	/
	2.3 The recycle materials	/
	2.4 The temperature system	2.4.1 Air conditioning
		2.4.2 Heating
3. Practice for green building	3.1 Concept	/
	3.2 3.3 Design	/
	3.3 Construction	/
	3.4 System	/
	3.5 Technology	/
4. Others	4.1 Summary	/
	4.2 Key date	/
	4.3 Others	/

After going through the context document , the preliminary case index made up of 4 elements are decided, including the project information ,result of practice ,practice of green building and others. Taken the availability of the information of green building experience, this paper carefully examines the context documents and chooses the final index shown in Table 1. According to the extent these elements can be broken down, this research resolves each element into different levels. The result of practice has two sub-levels while others only have one. Project information can be split into five more detailed information includes project name, project type, green building rating, project size and region. Result of practice consists of four detailed expected results involves the use of renewable energy, the greening condition, the recycle materials and the temperature system. More specifically, the use of renewable energy can be categorized into the use of power and water totally. Similarly, the temperature system is categorized into air condition system and heating system. The practice for green building consists of five specific practices includes concept, design, construction and equipment system. There is also some other information about the cases such as summery and key date, which are summarized as others. This was done manually by the researcher and each case was copied into a text file. The text mining was then carried out using *Leximancer* and *Nvivo*. This produced a preliminary result identifying some useful information. A world frequency query provided the detailed node information and the number of codes shown in Table 2. The 'number of the items coded' illustrates the frequency this item is mentioned in all the documents and the 'number of references coded' shows the frequency of mention in the case. For example, the node *power* is mentioned nine times in the documents of all the cases and there are six cases involving this node.

Table 2. Template for case indexing green building practices

Nodes	Expected result					
	Energy		Greening	Recycle	Temperature	
	Power	Water			Air conditioning	Heating
Number of items coded	9	7	17	5	6	11
Number of references coded	6	6	8	4	4	5

Nodes	Project information		Practice					
	Location	Project type	Concept	Construction	Design	System	Technology	Technology
Number of items coded	10	10	7	7	7	14	13	13
Number of coding references	10	10	5	6	4	3	5	5

After the preliminary text-mining analysis, *Nvivo* was used to analyze the relationship between nodes in terms of their similarity. The node frequency of the 10 cases is shown in Table 3. 'System', 'Greening' and 'Technology' are the three most frequently mentioned nodes, with the coverage of 34.27, 29.64 and 15.99 percent respectively. The Table also highlights the different focus of experiences between cases. For example, case 1 has a large amount of experience with the system (19.08%), while case 4 is more with sustainable technology (7.75%). Table 3 is therefore used to find the best-match cases. For example, if the new problem involves experience with the power system, the GBEM model firstly analyzes the project background and then the percentage coverage. In this example, case 4 is identified as the best case because it has the most experience of concept, design, construction and technology, which enables the new case to obtain relevant information to solve the problem.

At the end of study, a project map is drawn to analyze the relationships between all the nodes and cases, as shown in Fig. 3. From this, it can be deduced that the cases with more input arrows have the most experience, while those with more output arrows have more occurrence. In this case, case 1 has the most input, meaning that it contains the most experience, a conclusions that is consistent with the result of the coverage analysis. The system with the most output confirms the popularity of the node 'System'. Fig. 3 also provides a comprehensive perspective of the structure of all the nodes. 'Power' and 'Water', for example, are the two sub-nodes under 'Energy'. Similarly, 'Air conditioning' and 'Heating' are the two sub-nodes under the 'Temperature' node. Even though there is no direct link between the 'Energy' node and other components, their relationship is reflected through these sub-nodes.

According to the result, the case 1 should provide experience in terms of system and technology, case 2 turned out to provide the most useful experience in water saving. Case 3 has the most relevant information in greening and recycling system. Case 4 is the best reference case for power saving and construction. Case 6 provide the most useful experience for air conditioning and case 7 provides the best concept experience. Case 8 is the best referencing case for heating. Case 5, case 9 and case 10 are not reckoned as the first optional cases for the relevant experience mining. Therefore, it can be clearly seen how two cases are related to each other through the links among these nodes, which indicates the path for finding the best case.

Table 3. Percentage coverage of nodes by case

Code	Percentage coverage (%)										
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Total
Energy											
Power	1.35	/	/	4.53	1.94	/	/	2.03	1.25	4.02	15.12
Water	/	2.41	/	2.15	1.94	1.21	/	1.4	0.94	/	10.05
Greening	2.61	4.08	7.65	5.96	/	/	1.42	2.16	2.97	2.79	29.64
Recycle	/	/	4.17	1.07	/	/	0.89	/	/	1.55	7.68
Temperature	2.43	/	/	/	/	4.03	/	3.37	1.25	2.17	13.25
Air conditioning	1.44	/	/	/	/	3.3	/	2.86	/	2.17	9.77
Heating	/	/	/	/	1.82	4.19	/	5.47	1.25	2.17	14.9
Concept	/	/	0.7	1.19	1.71	/	4.35	/	/	0.63	8.58
Construction	/	1.86	/	4.05	1.48	0.97	/	0.7	1.41	/	10.47
Design	/	/	/	3.93	/	0.97	/	0.95	0.94	/	6.79
System	19.08	/	/	/	/	6.04	/	9.15	/	/	34.27
Technology	4.95	/	/	7.75	/	0.81	1.78	0.7	/	/	15.99
Total	35.01	14.47	12.52	33.61	14.71	25.95	9.59	31.08	14.85	16.43	/

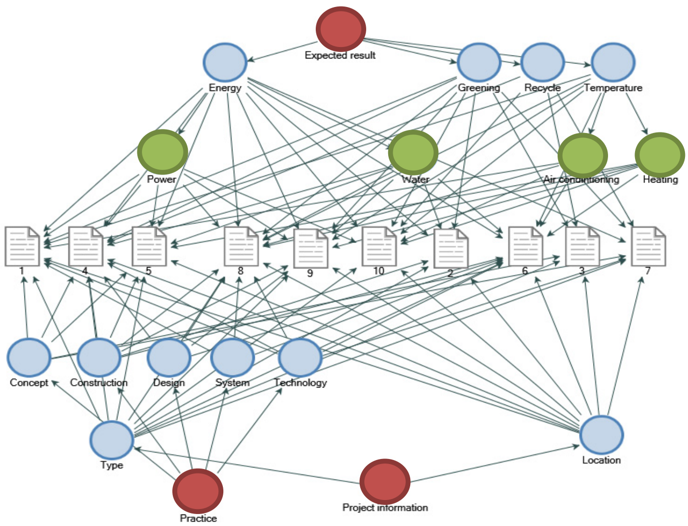

Fig.3. Case storage structure of the GBEM model

5. Conclusion and future research

This paper presents the GBEM model as an aid to solving green building problems by CBR together with a framework for green building experience mining. Text-mining techniques are used to enable the translation of the semantic questions involved into theoretical problems. A template is provided for representing the case experiences of green building and a database of 10 cases is established in a case study to demonstrate the use of the model.

With the increasing use of technology to collect, store and combine continuous data streams, there is a corresponding need to find ways of exploiting big data by such methods as data mining. Similarly, there is also a need to recapture hard-won knowledge from the past as an aid for future decision-making. In combining both approaches, the paper contributes to research into green building decision making in allowing large databases of existing successful green building practices to be combed for potential solutions to current problems. The breaks new ground in combining the CBR model with text analysis in green building decision making, broadening the domain of CBR application in terms of sematic problem solving and enabling the translation from an oral described problem into a theoretical problem with a certain answer.

The approaches used in this paper are intended to prompt more investigation into the use of experience mining in green building and beyond, and for increased attention to the organization of text information. Future research would benefit from the further empirical study of the application of text mining to transforming semantic descriptive questions into quantity problem.

References

[1] Eichholtz P, Kok N, Quigley JM. The Economics of Green Building. Review of Economics and Statistics. 2012;95(1):50-63.
[2] Glaeser EL, Kahn ME. The greenness of cities: carbon dioxide emissions and urban development. Journal of urban economics. 2010;67(3):404-18.
[3] Kahn ME. Urban growth and climate change. Annu Rev Resour Econ. 2009;1(1):333-50.
[4] Zheng S, Wang R, Glaeser EL, Kahn ME. The greenness of China: household carbon dioxide emissions and urban development. Journal of Economic Geography. 2010:lbq031.
[5] Yeh I-C. Quantity Estimating of Building with Logarithm-Neuron Networks. Journal of Construction Engineering and Management. 1998;124(5):374-80.
[6] Pal SK, Shiu SC. Foundations of soft case-based reasoning: John Wiley & Sons; 2004.
[7] Hinrichs TR. Problem-solving in open worlds: a case study in design. 1992.
[8] Ananiadou S. National centre for text mining: Introduction to tools for researchers. Retrieved February. 2008;8:2009.
[9] Yau NJ, Yang JB. Case-based reasoning in construction management. Computer-Aided Civil and Infrastructure Engineering. 1998;13(2):143-50.
[10] Hu X, Xia B, Skitmore M, Chen Q. The application of case-based reasoning in construction management research: An overview. Automation in Construction.
[11] An S-H, Kim G-H, Kang K-I. A case-based reasoning cost estimating model using experience by analytic hierarchy process. Building and Environment. 2007;42(7):2573-9.
[12] Chen D, Burrell P. Case-Based Reasoning System and Artificial Neural Networks: A Review. Neural Computing & Applications. 2001;10(3):264-76.
[13] Ng ST, Cheng KP, Skitmore M. A framework for evaluating the safety performance of construction contractors. Building and Environment. 2005;40(10):1347-55.
[14] Karshenas S. Predesign Cost Estimating Method for Multistory Buildings. Journal of Construction Engineering and Management. 1984;110(1):79-86.
[15] Yau N-J, Yang J-B. Applying case-based reasoning technique to retaining wall selection. Automation in construction. 1998;7(4):271-83.
[16] Quinlan JR. Induction of decision trees. Machine learning. 1986;1(1):81-106.
[17] Kibert CJ. Sustainable construction: green building design and delivery: John Wiley & Sons; 2008.
[18] Glavinich T. Contractor's guide to greenbuildingconstruction: management, project delivery, documentation, and risk reduction, ed. illustrated2008. USA: John Wiley and Sons.
[19] HU W-s, JIANG L-y, XUE M, YUAN N. The Practice of Sustainable Architecture in Campus——The Design of Green Demonstration Building of Logistical Engineering University [J]. Journal of Logistical Engineering University. 2010;1:001.
[20] Stankovic B, Kostic A, Popovic MJ. Analysis and comparison of lighting design criteria in green building certification systems —Guidelines for application in Serbian building practice. Energy for Sustainable Development. 2014;19:56-65.
[21] Azizi NSM, Fassman E, Wilkinson S, Ani AIC. Management Practice to Achieve Energy Efficiency Performance: Green versus Conventional Office Building in Malaysia. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction. 2013;5(4):205-14.
[22] Shen L-Y, Ochoa JJ, Zhang X, Yi P. Experience mining for decision making on implementing sustainable urbanization—An innovative approach. Automation in Construction. 2013;29:40-9.
[23] Chua D, Li D, Chan W. Case-based reasoning approach in bid decision making. Journal of Construction Engineering and Management. 2001;127(1):35-45.
[24] Hammond KJ, editor CHEF: A Model of Case-based Planning. AAAI; 1986.
[25] Ashley KD. Modeling legal arguments: Reasoning with cases and hypotheticals: MIT press; 1991.
[26] Watson I, Marir F. Case-based reasoning: A review. The knowledge engineering review. 1994;9(04):327-54.
[27] Romero C, Ventura S. Educational data mining: a review of the state of the art. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 2010;40(6):601-18.