Proactive problem-solver for construction

Wen-der Yu [a,*], Jyh-bin Yang [a], Judy C.R. Tseng [b], Shen-jung Liu [c,d], Ji-wei Wu [e]

[a] Inst. of Const. Management, Chung Hua University, Hsinchu, Taiwan, ROC
[b] Dept. of Computer Science and Information Eng., Chung Hua University, Hsinchu, Taiwan, ROC
[c] Dept. of Administration, CECI Engineering Consultants, Inc., Taiwan, ROC
[d] Dept. of Civil Engineering, National Taiwan University, Taipei, Taiwan, ROC
[e] Dept. of Computer Science, National Chiao Tung University, Hsinchu, Taiwan, ROC

ARTICLE INFO

ABSTRACT

Construction is an experience-based discipline. Knowledge or experience accumulated from previous projects plays a very important role in successful performance of new works. More and more construction organizations have adopted commercial Knowledge Management Systems (KMSs) to develop their own Knowledge Management (KM) functionalities. Most of the existing KMSs adopt Communities of Practice (CoPs) for knowledge sharing and exchange. Such an approach is found on the reactive problem-solver (RPS); that is, the problem raised by the questioner in the CoP has to "wait" for the "solution knower" to respond (or reply). Previous research indicated that the RPS approach may suffer in poor time and cost effectiveness. This paper proposes a Proactive Problem-Solver (PPS) approach for the problems encountered in construction engineering and management. Unlike RPS, the PPS proactively solves the problem based on lessons learned from previous projects. Should the solution be not available; the PPS dispatches the problem to the most appropriate domain experts so that the problem can be tackled timely and efficiently. A case A/E consulting firm is selected for implementation of the proposed PPS to demonstrate its applicability. It is shown that the proposed PPS improves more than 89.5% of efficiency both for timeliness and cost-saving of problem-solving. The proposed PPS demonstrates great potentials for improvement of emergent problem solving and enhancement of market competitiveness of a construction organization.

1. Introduction

Problem-solving is in the center of daily operations for construction organizations [1]. Since Construction Engineering is an experience-based discipline, knowledge accumulated from previous projects provides the key to solve similar problems encountered in future projects. Current practice of knowledge management system (KMS) has established an operational framework and a platform for problem solving in construction engineering and management [2]. One most commonly adopted problem-solving platform in a KMS is the Communities of Practice (CoP). According to Wenger and Snyder (2000), the CoP was defined as a group of people informally bound together by shared expertise and passion for a joint enterprise [3]. In an engineering consulting firm, a CoP is usually implemented as a subsystem of a KMS, which forms a virtual community for a group of people (the members of the CoP) who share interests on a professional/technical subject, e.g., structural design, geotechnical issues, material specification, contract management, etc. The common KM activities of the members in a CoP include [4]: (1) publishing articles for requesting of information on the electric forum system; (2) responding to the published articles by publishing additional articles; (3) holding meetings for members to build sense of belonging.

The KMS approach for problem solving poses several desirable features over other methods (such as Systems Engineering) including: (1) the experienced-based solutions that were implemented and verified in real world cases are more realistic and practical than theoretical solutions generated by analytic methods; (2) the collective intelligence supported by domain experts in the CoP provides a broader knowledge base and diverse perspectives to generate a more effective solution; (3) the KMS records all discussions while deriving the solution in CoP, so that the "experiences" of problem solving are automatically stored for future use.

Although the KMS approach poses many desirable features for construction problem solving, there are also essential drawbacks that exist in the traditional KMSs. The most critical disadvantage of a KMS for solving emergent problems is its nature of "reactive mode" of KM (referred hereafter as Reactive Problem-Solver or RPS). That is, the problem raised by the questioner has to wait (passively) for replies and responses from the "solution knower" in the CoP of the KMS. Previous research has indicated that such approach can be the bottleneck to improve the performance of the KMS due to poor time

* Corresponding author. Tel.: +886 3 5186748; fax: +886 3 5370517.
E-mail address: wenderyu@chu.edu.tw (W.D. Yu).

and cost effectiveness of the RPS [2]. Moreover, the verification, storage and retrieval of previous solutions (also called "lessons-learned") cause difficulties for the successful application of the traditional KMS for construction problem solving.

The present paper aims at addressing the abovementioned problems encountered in the traditional KMS for construction problem solving. A proactive problem-solving system, namely Proactive Problem-Solver (PPS), is proposed to improve the disadvantages of the traditional RPS. In contrast to the traditional RPS, the proposed PPS proactively "tackles" the problem posed by the questioner in a CoP and replies with the most appropriate solution based on previous lessons-learned. Should the solution be unavailable from historic lessons, the PPS "proactively" dispatches the problem to the most appropriate domain experts (in the organization) who are knowledgeable of relevant tacit (implicit) knowledge and solve the problem manually.

The rest of the paper starts with reviews of related works to provide required backgrounds for PPS; the model of PPS is then proposed with detail discussions of the required functions and components; then a case study is conducted to develop and test a web-based implementation of the proposed PPS for a leading A/E consulting firm in Taiwan; discussions on system strengths, limitations, and potential applications are addressed based on observations from case study results; finally, conclusions and recommendations are provided to interested readers.

2. Review of related works

The term "Proactive Problem-Solving" is not found in literature. However, related issues and similar functions of PPS addressed in the problem statement can be found in some existing works.

2.1. Problem-solving in construction

Problem solving plays the central role of daily construction operations. Li and Love [1] developed a framework of problem-solving for construction engineering and management. Their research identified several characteristics of construction problems that should be tackled in order to solve them quickly, correctly, and cost-effectively, such as the ill-structure nature, inadequate vocabulary, little generalization and conceptualization, temporary multi-organization, uniqueness of problems, and hardness in reaching the optimal solution. Two areas of problem-solving researches tackle the above-mentioned issues: the cognitive science and decision support system (DSS). The cognitive science-based approach is the most widely adopted as it is the basis for manual problem-solving techniques. The decision support systems (DSSs) are widely tested in academia. Many researchers develop their own DSSs for special purposes, such as cost estimation, technology selection, mark-up decision-making, duration estimation, etc.

In addition to these two areas, Yu et al. [2] propose a third approach called Knowledge Management integrated Problem-Solver (KMiPS) to solve emergent construction problems. The KMiPS adopts a KMS and a special designed CoP, namely SOS, for emergent problem solving. Yu et al. proved that the KMiPS achieved both quantitative and qualitative benefits better than the traditional problem-solving approaches. Their research showed that KMS provides desirable functions to tackle the special characteristics of construction problems identified by Li and Love. However, some essential drawbacks (such as "reactive mode" of problem solving) exist in the traditional KMSs, which may cause poor performance of timeliness and cost effectiveness.

2.2. Knowledge classification and knowledge map

While applying KMS for construction problem solving, the storage and retrieval of previous lessons-learned are crucial. Such issues become critical as the number of historic lessons grows. As a result, the methods of knowledge classification or knowledge map were developed. Kim et al. [5] proposed a practical method for capturing and representing the knowledge that is critical in knowledge management. The method employs a knowledge map as a tool to represent the knowledge of a firm. Their procedure consists of six steps: (1) defining organizational knowledge; (2) analyzing process map; (3) extracting knowledge; (4) profiling knowledge; (5) linking knowledge; and (6) validating map knowledge. Effective knowledge maps help identify intellectual capital, socialize new members, enhance organizational learning, and anticipate impending threats and/or opportunities [6]. Caldas et al. [7] proposed an automatic document classification method based on text mining. Their work successfully classified 4000 documents automatically with the Construction Document Classification System (CDCS) they developed. Although the abovementioned methods provide feasible alternatives for knowledge classification of the previously accumulated knowledge, none of them addresses the consideration of business domains and the organizational structure of the firm that may significantly affect the effectiveness of classification of the knowledge for problem solving.

2.3. Automatic Answering System (AAS)

Automatic Answering System (AAS) serves as a domain expert who is able to answer the question posed by the questioner instantly. Various types of AAS's have been developed in construction industry. The Advanced Construction Technology System (ACTS) was developed in the University of Michigan at Ann Arbor by Ioannou et al. [8]. ACTS provides a technology information system for construction planners and managers to select the most appropriate state-of-the-art construction technologies during the project planning stage. More than 400 technologies are recorded with 25 attributes such as general description, cost benefit, construction constraints, special application, operation environment, test criteria, etc. The Architecture and Engineering Performance Information Center (AEPIC) was developed by Loss at the University of Maryland [9]. The AEPIC provides information of failures so that the mistakes won't be repeated again. The On-Line Reference Library (OLRL) was developed by the Bechtel Inc. to provide engineers with real-time reference manuals of SPECs. The Civil Engineering Information System (CEIS) of Kajima Corp. is similar to ACTS and OLRL, which stores more than 300,000 technical documents [10]. Even though the abovementioned systems provide some features of AAS, most of them are database systems equipped with search functions. None of them provides complete functionalities required for proactive problem solving, such as automatic problem characterizing, intelligent information retrieval, problem dispatching, and solution repository. Moreover, they are information system rather than problem-solving system.

2.4. Lessons-learned system

Another issue related to construction problem-solving is the compilation of previous learned knowledge that is useful to solve future problems. Such knowledge is usually called "lessons-learned". There have been many existing lessons-learned systems reported in literature, which provides references for the present research. The Hypermedia Constructability System (HCS) was developed in collaboration between the Indiana Department of Transportation (INDOT) and the Purdue University [11]. The HCS stores historic lessons-learned in multi-media format so that construction engineers can learn from previous lessons more effectively. The Constructability Lessons Learned Database (CLLD) and Integrated Knowledge-Intensive System (IKIS) were developed by Kartam and Flood [10,12] to provide a repository for previously learned lessons. The major difference between CLLD and IKIS and the abovementioned lessons-learned systems is that the former verifies historic lessons-learned by the domain experts before storing in the database.

A recent work by Mohamed and AbouRizk developed a knowledge representation schema for construction problem solutions (lessons-learned) [14] based on the theory of inventive problem-solving (TRIZ) [15]. Their schema consists of three major components: (1) the main functions/effects of the solution; (2) the contradiction set of the encountered problem; (3) the resolution principle that best represents the solution. Mohamed and AbouRizk also developed a computer system to implement the proposed schema. Their method provides a framework for efficient knowledge representation for construction lessons-learned. One weakness of the schema is that only principles but no details of problem resolution lessons are stored, which may cause difficulty of users to reapply the lessons-learned.

The Construction Industry Institute (CII) developed a Lessons-Learned Wizard (LLW) with the package of constructability program [13]. The LLW is a computer aided information system that helps the engineers to record and retrieve the lessons learned from historic projects. The major components of a lesson-learned file (LLF) captured by LLW consists of: (1) problem description — describing the problem encountered in construction process; (2) information of the LLF writer and approvers — providing contacting information for further consultation; (3) solution description — describing the technical and procedural details of problem resolution; (4) evaluation of solution — assessment of the effectiveness and benefits resulted from the lesson-learned. Compared with the other methods mentioned above, the CII's LLF is more suitable for construction problem-solving due to two reasons: (1) more technical and procedural details are provided so that it is easier for users to reapply the LLFs; (2) the LLFs are verified and assessed before compilation, so that the solution stored in the LLF is more reliable and practical.

3. Proposed Proactive Problem-Solver (PPS)

3.1. Description of required functions

According to the problem statement and literature reviews, several functionalities are identified for the proposed PPS, including: (1) a knowledge classifications scheme (namely Knowledge Map or KMap) that appropriately represents the lessons-learned of an organization and accurately defines problem encountered; (2) a descriptive scheme for the expertise of the domain experts (namely Expert Map or EMap), which properly reflects knowledge (expertise) of the domain experts accumulated previously; (3) a set of algorithms for retrieval of the most relevant lessons-learned in repository; (4) a problem dispatching mechanism for diverting the posed problem to the most appropriate domain experts; and (5) a repository of lesson-learned files (LLFs) from previous projects and the required functions of lesson learning.

Four modules are proposed to fulfill the functionalities identified above: (1) the Knowledge/Expert Map (K/EMap) — providing classification and descriptive schemes for the knowledge in the domain and the expertise of the relevant domain experts; (2)

Automatic Problem Answering (APA) module — solving the posed questions automatically based on the historic LLFs; (3) Automatic Problem Dispatching (APD) module — dispatching the posed problem to the most appropriate domain expert(s) when the problem is unsolved by APA; (4) Lessons-Learned Wizard (LLW) — accumulating historic LLFs based on the classification scheme of K/EMap in a LLF repository.

3.2. Knowledge/Expert Map (K/EMap)

The proposed PPS is kernelled with the knowledge and the experts holding the knowledge. In PPS, the domain knowledge is classified by Knowledge Map (KMap) while the domain experts are characterized by Expert Map (EMap). The KMap and EMap provide the ontology for modeling the knowledge and human assets of a construction organization.

A Multi-dimensional Knowledge Ontology (MKO) scheme is adopted for construction of the KMap. The multiple dimensions of MKO are represented with a vector of numeric codes. Fig. 1 shows the schematic representation of MKO. The MKO consists of three dimensions including: "Lifecycle code", "Product code", and "Technical code".

The MKO scheme represents a knowledge documents (e.g., LLF) with a vector of codes including: (1) Lifecycle code — describing the time reference of the knowledge including the different phases of a project lifecycle: feasibility analysis, comprehensive planning, basic design, detail design, construction/installation, testing, operation/maintenance, recycle, change, etc.; (2) Product code — describing the service/product related to the knowledge, such as bidding proposal, execution plan, QA plan, procurement plan, SPEC, alternative report, inspection report, calculation report, drawing, etc.; (3) Technical code — the technical classification of the knowledge, such as administration, human resource, civil (excavation, refill, site preparation, piping, etc.), structural (RC, PC, SS, underground construction, retaining structure, etc.), architecture (urban planning, building design, interior design, landscape, model, finishing, etc.), geotechnic (site investigation, pile, foundation, drainage, rock, stability, etc.), survey, highway, transportation, logistics, airport, hydraulics, harbor, material testing, etc. An example vector "(13, 20, 35)" can be interpreted as "the *design criteria* (product code = 13) for *segmental precast bridge construction* (technical code = 20) in *basic design* phase (lifecycle code = 35)".

The MKO model is adopted in light of: (1) effective and efficient classification of LLFs — MKO does not only provide a coding system for classification of knowledge documents, it also offers a hierarchical framework for association of relevant knowledge; (2) correct and accurate characterization of the problem — with MKO, the temporal dimension, product type, and detailed technical field of the problem and the associated solution (LLF) are accurately characterized, it makes future retrievals of historical lesson-learned more easily; (3) a basis for construction of the EMap — the domain expert is represented by the knowledge he/she holds. Therefore, MKO also provides a key attribute for characterization of the domain experts in the EMap.
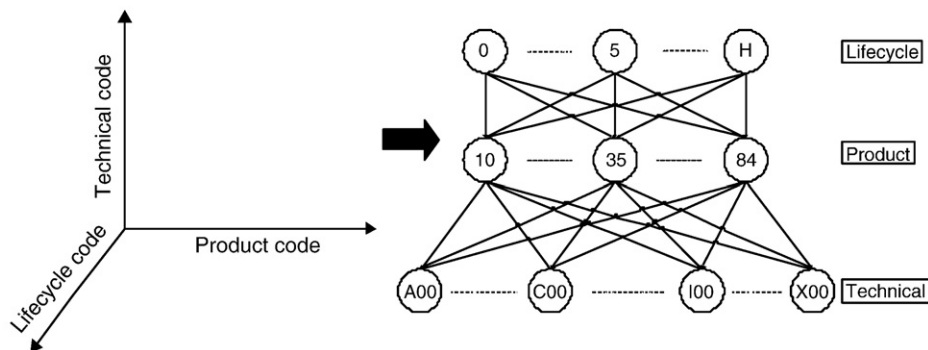

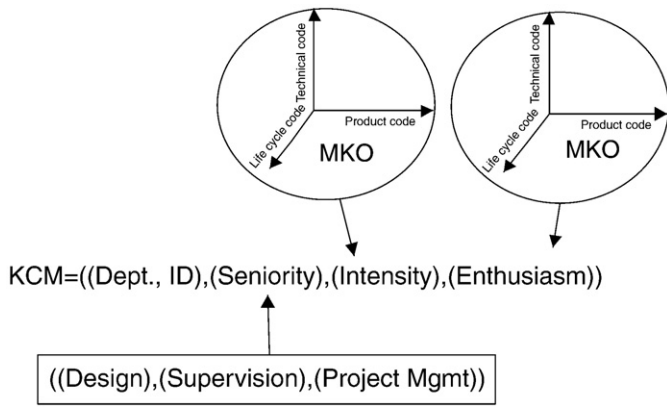**Fig. 1.** Representation of Multi-dimensional Knowledge Ontology (MKO).

Fig. 2. Representation of Knowledge Capacity Matrix (KCM).

Similar to KMap, a Knowledge Capacity Matrix (KCM) is proposed for representing the expertise of the domain experts. The KCM describes the knowledge capacity of a domain expert with a row vector containing three dimensions: (1) Seniority — recording the professional seniority (in years) of the expert, which reflects how experienced the domain expert is in a specific technical area; (2) Intensity — recording the intensity of work (work hours/seniority years) of a specific technical domain, which reflects the strength of expertise of the expert in a specific domain; (3) Enthusiasm — recording the historic performance of the expert in participation of KM activities automatically provided by the KMS, which reflects what technical domain the expert is especially interested in. The second (Intensity) and the third (Enthusiasm) dimensions are characterized with the MKO described above, so that the specific technical domain of knowledge that the domain expert associated with is identified. The KCM adopts a nested row-vector representation scheme. An example of KCM is shown in Fig. 2. In Fig. 2, the Seniority dimension consists of the experience records (in years) of three fields: Design, (Construction) Supervision, and Project Management. The other two dimensions (Intensity and Enthusiasm) are represented with the values of "expertise intensity" and "the knowledge value adding scores" (calculated by the method of KVAM [4]) that are characterized with MKO.

With the schemes of KMO and KCM, a domain expert is characterized according to his/her experiences, not only by a specific technical field but also by how experienced and intensively he/she was involved and how enthusiastic he/she is in that area.

3.3. Automatic Problem Answering (APA) module

The Automatic Problem Answering module (APA) is an Automatic Answering System (AAS) that searches the solution database to retrieve the most appropriate answers for the questioner. Several requirements are expected for an ideal AAS [17]: (1) tolerance of simple errors; (2) embodiment of some degree of "common sense"; (3) a relatively large and complete vocabulary for the subject matter to be treated; (4) acceptance of wide range grammatical constructions; (5) capability of dealing sensibly with partially understood input; and (6) providing the information and computations requested by the user. Previous research has developed an AAS for internet service [18] and tutoring assistant [19]. The conceptual design of APA module is planned as shown in Fig. 3 based on the AAS proposed by Wu et al. [19].

In Fig. 3, a question is posed by the questioner. The characteristics of the posed problem are analyzed by the APA. Then, APA searches the repository of historic lesson-learned files (LLFs) to find the most relevant LLFs. Finally, the most appropriate solution is retrieved from the relevant LLFs. The underlying algorithm for APA is described below:

1. Transforming problem description of the LLF into a characteristic vector (CV). In this step, the linguistic problem description of the LLF is transformed into a CV using the Vector Space Model (VSM) [20,21]. At first, the problem description is segmented into semantic fragments (terms) according to the domain keywords. In the present research, the Chinese Knowledge Information Processing (CKIP) database provided by the Institute of Information of the Academia Sinica, Taiwan is adopted for the segmentation of the document [22]. Second, the importance weightings of the CV associated with the keywords (terms) are calculated using the Importance Factor (IMF) method proposed by Wu et al. [19] as described in Eq. (1):

$$IMF_{ij} = \frac{L_j}{L_{i,\max}} \left(0.5 + 0.5 \frac{TF_{ij}}{TF_{i,\max}} \right) IDF_j, IDF_j = \log \left(\frac{N}{\sum_{i=1}^{N} C_{ij}} \right) \quad (1)$$

where IMF_{ij} is the weighting of the jth term in the ith question (Q_i) of the LLF repository; L_j is the length of the jth term; $L_{i,\max}$ is the maximum length of terms in Q_i; TF_{ij} is the number of occurrences for the jth term in Q_i; $TF_{i,\max}$ is the number of occurrences for the
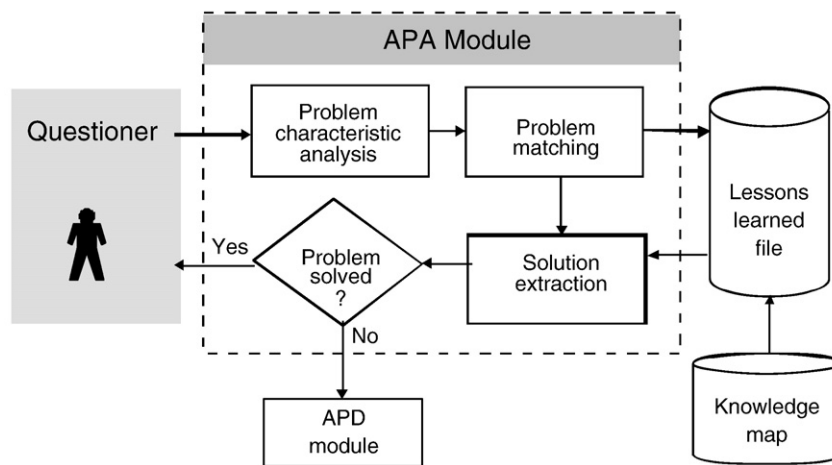

Fig. 3. Conceptual design of APA module.

term which has the maximum number of occurrences in Q_i; N is the total number of LLFs in the repository; and $C_{ij} = 1$ if Q_i contains term j and $C_{ij} = 0$, otherwise.

The CV of question Q_i can be represented with Eq. (2):

$$CV(Q_i) = \left\{ (k_1, w_{i1}), (k_2, w_{i2}), ..., \left(k_j, w_{ij}\right), ..., (k_n, w_{in}) \right\} \qquad (2)$$

where $CV(Q_i)$ is the CV of the ith question; K_j represents jth keyword; and W_{ij} is the weighting value of K_j in question Q_i.

2. Once a new question Q is posed by the questioner, the *problem characteristic analysis module* will transform the question Q into CV (Q) using the similar method of Step 1.

3. After problem characteristic analysis, the *problem matching module* will search the LLF repository to find the most relevant LLFs. The CV of the posed question is compared with all historic CV's stored in the LLF repository using the *inner product similarity measure* described in Eq. (3):

$$S_i = \sum_{j=1}^{n} \left(W_j \times W_{ij} \right) \qquad (3)$$

where S_i is the similarity between question Q and the problem description of the ith LLF; W_j is the weighting of the jth element of the $CV(Q)$; W_{ij} is the weighting of the jth element of the problem description of the ith LLF, i.e., $CV(Q_i)$.

4. The LLFs with higher similarity are considered more relevant to the question Q posed by the questioner, thus they are recommended as the solutions associated with the posed question.

If one or more solutions are found (according to a predetermined similarity threshold), they are retrieved by the *solution extraction module* from the LLF repository and replied to the questioner. However, if no solution is found, the problem remains unsolved and is diverted to the APD module.

3.4. Automatic Problem Dispatching (APD) module

Fig. 4 shows the conceptual design of the APD module. In Fig. 4, the unsolved problem is posted in the CoP of the KMS as an emergent problem and diverted simultaneously to APD module at the same time. In APD module, the problem characteristics is analyzed and used to find the most appropriate domain experts based on the EMap
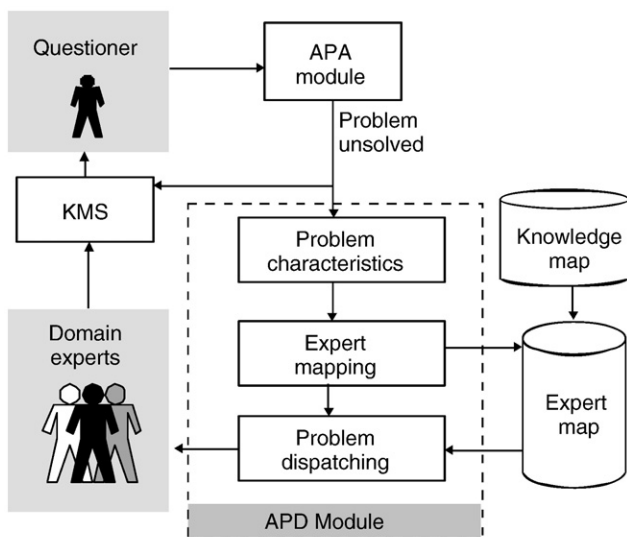

Fig. 4. Conceptual planning of APD module.

described previously. Then, the problem is dispatched to the most relevant domain experts for the possible solutions. Finally, the experts respond to the problem in a special CoP (in the case study, the special CoP is called "SOS") of the KMS.

To find the most appropriate domain experts, APD chooses from EMap the experts who match the problem characteristics as candidates first. Then a matching score is calculated for each candidate of domain expert with respect to the posed question using Eq. (4):

$$MS_i = W_1 \times S_i + W_2 \times I_i + W_3 \times E_i \qquad (4)$$

where MS_i is the overall matching score of the ith domain expert with respect to the posed question; S_i is the *Seniority* score of the ith domain expert; I_i is the *Intensity* score of the ith domain expert; E_i is the *Enthusiasm* score of the ith domain expert; W_1, W_2 and W_3 are weightings for the Seniority score, Intensity score, and Enthusiasm score, respectively; noted that the sum of W_1, W_2 and W_3 should equal to 1.

In Eq. (4), the weightings W_1, W_2 and W_3 are determined by the questioner based on his/her understanding of the application domain. For example, if the application relies more on experience, then the weighting of *Seniority* is emphasized, and so forth.

The *Seniority* score in Eq. (4) is further calculated with Eq. (5):

$$S_i = W_{11} \times DE_i + W_{12} \times SE_i + W_{13} \times PME_i \qquad (5)$$

where DE_i is the seniority of the design experience for the ith domain expert measured in years; SE_i is the seniority of the supervision experience for the ith domain expert (in years); PME_i is the seniority of the project management experience for the ith domain expert (in years); W_{11}, W_{12} and W_{13} are weightings for Design Experiences (DE), Supervision Experiences (SE) and PM Experiences (PME), respectively.

Similar to Eq. (4), the questioner can determine the weightings (W_{11}, W_{12} and W_{13}) arbitrarily according to his/her understanding of the application domain and the sum of W_{11}, W_{12} and W_{13} should be 1.

3.5. Lessons-Learned Wizard (LLW)

The LLF repository required in APA module is constructed by a Lessons-Learned Wizard (LLW) proposed by the Construction Industry Institute (CII) [13]. The LLW captures lessons-learned right after a problem is solved in the KMS. The LLW is integrated with an internet questionnaire surveying system that allows the questioner to evaluate the solution he/she obtains. The LLF associated with the problem contains the following information: (1) the subject of the problem — the topic of the problem; (2) the description of the problem — detailed descriptions of the posed question; (3) the questioner — the name and department of questioner; (4) the solution — detailed descriptions of the suggested solution; (5) the responder — the names and departments of the responders who provided the suggested solutions; (6) attachments — the supplementary materials for the solution; (7) benefits evaluation — assessments of all benefits resulted by the solution including time, cost, quality, technical improvement, regulation impacts, etc. The LLFs stored in the LLF repository are classified with the MKO scheme of KMap.

3.6. Integrated Model of Proactive Problem Solving (MPPS)

The integrated framework of MPPS is depicted in Fig. 5 in which, MPPS solves construction problems in two modes: (1) Automatic problem answering mode (APA mode) — the problem-solving process is shown in Fig. 5 as bold solid arrows, where the solution is searched automatically from LLF repository according to the problem characteristics; (2) Automatic problem dispatching (APD) mode — the problem-solving process is shown in Fig. 5 as dashed arrows, where
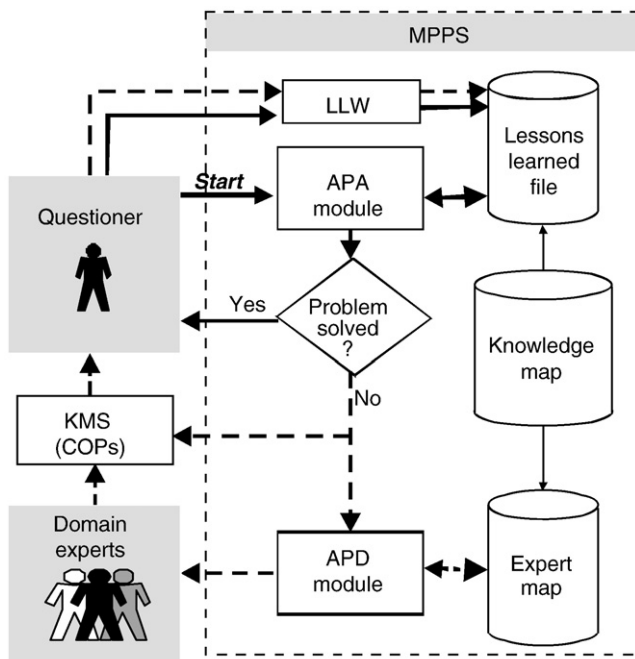
Fig. 5. Integrated framework of MPPS.

the unsolved problems (by APA mode) is automatically dispatched to the most relevant domain experts according to the problem characteristics and the Knowledge Capacity Matrix (KCM). The functions of problem solving in the traditional KMS is preserved and exercised in MPPS as shown in Fig. 5 where the unsolved problem is posted in the CoPs of the KMS before entering the APD mode.

Both the problems solved by APA and APD modes are considered as new lessons learned for future problems. This process is actually a verification of the knowledge to generate a higher level of intellectual asset called "wisdom" [16]. This verification process is performed by LLW as shown in Fig. 5.

4. System implementation

The proposed MPPS has been implemented in the KMS of a leading A/E consulting firm in Taiwan, CECI. The KMS of CECI is developed based on a commercial platform — Microsoft SharePoint®. However, the original software has been customized to fit in the specific requirements of the firm. One of the major customizations is the specialized emergent problem-solving system, called SOS, to provide real-time aids for engineers/managers who are encountered with emergent problems [2]. The SOS system has been proved to be very beneficial to the firm. Both tangible and intangible benefits were resulted significantly [23].

4.1. Problems faced the existing system

The SOS system is a specialized CoP that includes all staffs of the firm as its members. Once a question is posed by a questioner in SOS, it will prompts automatically on the portal page of the KMS for all members. As described in Introduction, the essential problem of the existing CoP in solving emergent problems is that the posed problem needs to wait (passively) for replies and responses from the "solution knower" to provide his/her solutions. Such "passive" mode of problem solving assumes that the domain experts can "see" the problem and respond with their solution timely. However, previous research found that such RPS approach has caused inefficiency of timeliness and cost-

effectiveness of the KMS [2]. A proactive problem-solving approach should be developed.

4.2. System implementation

The proposed MPPS has been implemented with the SOS system of CECI. A Proactive Problem-Solver (PPS) is developed and tested. The prototype PPS consists of all four required elements of MPPS: (1) K/EMap — a knowledge map (KMap) is constructed based on the MKO scheme and an Expert Map (EMap) is constructed based on KCM scheme; (2) APA module — the APA module is developed to perform the APA problem-solving mode (APA mode); (3) APD module — the APD module is developed to perform the APD problem-solving mode (APD mode); (4) LLW — the LLF repository is established consisting of 908 historic problem-solving LLFs of SOS system accumulated in the last three years.

5. System testing and performance evaluation

In order to verify the proposed PPS, system evaluation experiments are designed and conducted. The experiments consist of two parts: (1) effectiveness test — testing the validity of MPPS and correctness of PPS in finding the solutions for the posed questions; (2) efficiency test — testing the timeliness and cost-saving performance of PPS compared with the traditional approaches.

5.1. Data collection and experiment design

The testing data were collected from real world emergent problem-solving cases of the case A/E consulting firm, CECI, from 2005/01 to 2009/08. Totally 908 historical cases and the associated LLFs were collected. The major problem categories and their associated percentages of the 908 cases are: Architectural (14.76%), Civil (13.51%), Structural (13.51%), Geotechnic (10.04%), Electrical (7.16%), Railway (5.76%), Environmental (4.80%), Mechanical (4.72%), Highway (4.43%), Materials (3.47%), Hydraulic (2.58%), and Others (15.26%).

Assume that every problem-solving case has one "correct" (most relevant) solution, which is stored in the LLF repository. The experiments are designed to test PPS with two sets of data: (1) Original Set — testing the capability of PPS to retrieve the exact solution for a specific question with the original question description (same as that of the LLF); and (2) Similar Set — testing the capability of PPS to retrieve the most relevant solution for a specific question with the modified Similar (but articulated differently) question description from the original one of the LLF. If the PPS is able to retrieve the correct solution for both sets of testing data, the PPS system is verified and the proposed MPPS is validated. In order to generate the testing data for the Similar Set, sixty-three mangers/senior engineers of the case A/E consulting firm were asked to play the role of the domain experts. The problem descriptions of the 908 historic LLFs were presented to the 63 domain experts (with domain experiences for ten to twenty-five years from all related technical departments) who were then requested to provide 1 to 3 similar but articulated differently question descriptions for testing of PPS. After collection, 1368 question descriptions are generated for the Similar Set based on the 908 original questions. However, it was found that 64 question descriptions provided by the domain experts was invalid after reviewing the generated data sets. Those data sets were excluded from the testing set. The rest 1304 Similar Set and the 908 Original Set were used to test the effectiveness and efficiency of PPS. It is noted that the historic lesson-learned file (LLF) only provides the "preliminary solution", and the user of the LLF has to develop his/her own final solution. Therefore the performance evaluation is this section is conducted only on the efficiency of finding "preliminary solution".

5.2. Testing of effectiveness

There are two types of testing for the effectiveness of information retrieval systems: (1) *Precision* — measurement of the effectiveness of a retrieval system to retrieve only the relevant answers; and (2) *Recall* — measurement of the effectiveness of a retrieval system to retrieve all the relevant answers. Since the *Recall* is more important than *Precision* in the information retrieval system such as emergent problem solver [24], only *Recall* is adopted for effectiveness testing of PPS. The index of *Recall* is defined in Eq. (6) [24]:

$$Recall = \frac{Number\ of\ relevant\ answers\ retrieved}{Total\ number\ of\ relevant\ answers} \qquad (6)$$

where the numerator is the total number of relevant LLFs retrieved by PPS, and the denominator is the total number of all relevant LLFs stored in the LLF repository.

In the present research, it is assumed that there is only one relevant (correct) LLF for each question. Let R_n be the probability that APA is able to retrieve the relevant solution for the posed question with n retrieved (recommended) LLFs, i.e. the *Recall*, the R_n is defined as Eq. (7):

$$R_n = \sum_{i=1}^{N} \frac{R_{ni}}{N}, \begin{cases} R_{ni} = 0, & \text{if the retrieved LLFs does not include the correct answer.} \\ R_{ni} = 1, & \text{if the retrieved LLFs includes the correct answer.} \end{cases} \qquad (7)$$

where N is the total number of testing questions; n is the number of retrieved LLFs; R_{ni} is a true/false testing value ("1" for true, "0" for false) of whether the retrieved LLFs include the correct answer.

Both of the Original Set and the Similar Set are used for testing. The parameter n is an integer variable varied from 1 (retrieve only the LLF with highest similarity) to 10 (retrieve top ten LLFs with the highest similarities). The testing results are shown in Fig. 6.

It is found from Fig. 6 that the *Recall* of the Original Set is very close to 1 (100% correct) as $n \geq 2$, while the *Recall* of the Similar Set is nearly 90% after as $n = 10$. As a result, the effectiveness of PPS is generally verified.

5.3. Testing of efficiency

The two indexes of efficiency for PPS are the timeliness and cost-saving of the system. According to a previous research, the average cost for solving (finding the preliminary solution) a single problem by the traditional KMS is TWD 4075 (roughly equals to USD 123.5) [23]. With the aid of PPS, the cost for retrieving the relevant LLF (if it exists in the LLF repository) is almost costless. The improvement of cost-saving is about 99.99%. Thus, the cost-saving efficiency of PPS over the traditional approach is verified.

In regard to the timeliness of PPS, three components of time required for problem-solving with PPS are identified:

(1) Processing time of PPS (S)
Assume that S is the execution time required to search the relevant LLFs with PPS. It is considered constant for all questions. By monitoring the execution time of PPS, the average duration of S is about 3.2 s.
(2) The processing time of the questioner (F)
The time required for the questioner to find out the really relevant LLF from all LLFs retrieved by PPS. Assume F seconds are required to process one LLF, then processing n LLFs requires $F \times n$ seconds. In the present research, F is assumed to be 10 s.
(3) The time required to generate a solution manually (p)
As the APA is unable to retrieve the relevant solution (it may be due to that there is no relevant LLF available or PPS is unable to identify the relevant LLF), the solution needs to be generated manually by the domain experts. Assume that it takes time p to generate a solution manually. According to a previous research [23], the average time required to generate the solution for a posed question is 2.68 days. With the help of APD, the time required could be less. This time is required only when there is no relevant LLF found by PPS. Thus, the time required to generate solution manually can be calculated by $(1-R_n) \times p$.

Based on the above analysis, the time required to generate a solution with PPS can be calculated with Eq. (8):

$$T = S + F \times n + (1-R_n) \times p \qquad (8)$$

where T is the measure of timeliness of problem-solving; S, F, n, and R_n are defined previously.

Based on Eq. (8), the timeliness of PPS for the two sets of testing data are compared and shown in Fig. 7. It is found that the average time required to solve the Original Set is 0.887 min as $n = 5$. It increases slightly as n increases due to the time required to single out the relevant LLFs. The average time required to solve Similar Set is much higher. This is due to the time required for manual problem-solving (p). However, the problem-solving time of PPS is about 400 min (6.67 h) as $n > 9$. It is much lower than the traditional KMS (which required 2.68 days = 3859.2 min). The timeliness improvement is about 89.6%. Thus, the timeliness efficiency of PPS is also verified.
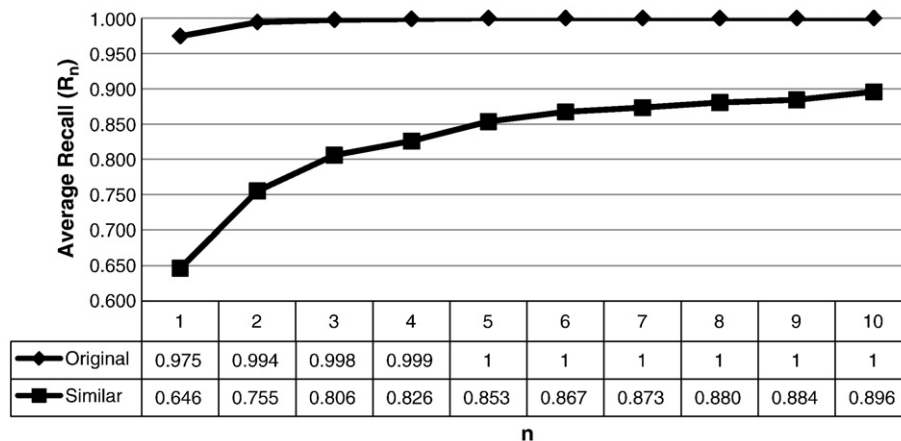

	1	2	3	4	5	6	7	8	9	10
Original	0.975	0.994	0.998	0.999	1	1	1	1	1	1
Similar	0.646	0.755	0.806	0.826	0.853	0.867	0.873	0.880	0.884	0.896

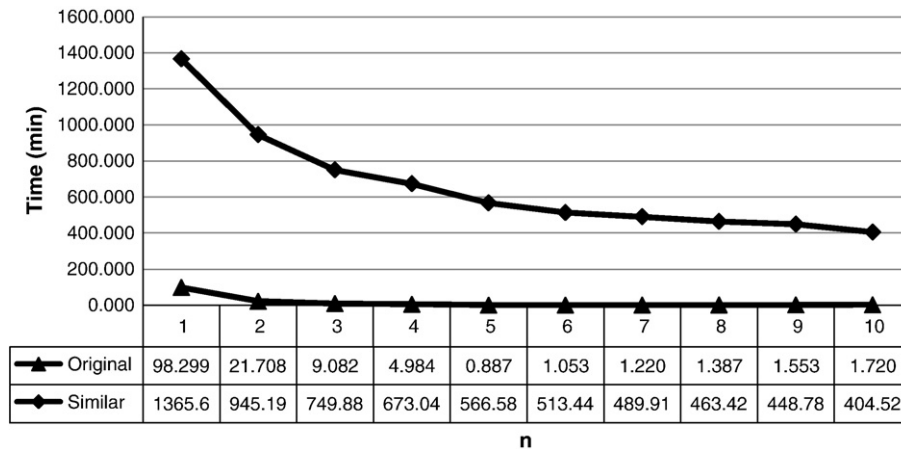**Fig. 6.** Comparison of recall (Original vs. Similar Set).

Fig. 7. Comparison of timeliness (Original vs. Similar set).

6. Discussions

The proposed PPS provides construction engineers and managers a different approach of problem-solving. This section addresses the strengths, potential extensions, and limitations of such a method.

6.1. Strengths of PPS over traditional problem-solving approaches

The primary strength of the proposed PPS over the traditional KMS resides in its proactive mode of problem-solving. Instead of "waiting for" solution, the proposed PPS take initiative to retrieve relevant solutions from historic LLFs and identify the domain experts who are most capable of resolving the posed question. Based on the results of case study, the benefit of time-saving is 99.99% for the Original Set and 89.5% for the Similar Set. Moreover, the problem-solving is almost costless if the relevant historic LLFs exist. Nevertheless, the improvement of timeliness in problem-solving by PPS should enhance the competitiveness of the firm in the market since it equips the firm with a powerful tool to react with all kinds of emergent issues that happen in the firm's daily business operations. Any of the issues can become a trouble source of increased costs or time delays if it is not tackled properly.

The second advantage of the proposed PPS over the traditional KMS is that it facilitates the intellectualization of the study firm's tacit and explicit knowledge assets. Refer to Fig. 5, the historic lessons-learned for solving construction problems need to be compiled into LLFs; KMap and EMap are then constructed to visualize the firm's intellectual assets. The top management can plan the firm's knowledge management strategies more effectively based on the profile of the firm's intellectual assets. Moreover, the proposed PPS provides a mechanism to incorporate the tacit knowledge of all staffs of the firm to generate solutions for the posed questions. Such mechanism offers a perfect environment for the Nonaka's knowledge creation spiral [25] so that Medici's Effect [26] can take place. As a result, the knowledge creation of the firm is expedited and more fruitful.

6.2. Potential extensions and applications of PPS

The proposed PPS has demonstrated its potentials in emergent problem solving for construction. Some future extensions may be pursued. An application of PPS is to help engineers in preparation of the proposals. A critical component (may be the most important one) of the proposal is "Critical issues analysis". The PPS can provide solutions to those issues addressed in the request for proposal (RFP).

A second extension of the current research is to integrate PPS with the ubiquitous technology to enhance the real-time problem solving.

For example, the wireless communication technology provides possibility of accessing PPS anytime anywhere. Such technology is very desirable especially for emergent problem solving on construction site. The integration of the two technologies will enable the site engineers/managers to fully utilize the advantages of PPS.

The methodology developed in the current research can also be applied to problem-solving of the other domains. For example, the legal affairs and management consulting are two promising areas for application of the proposed PPS. Extensions to other fields such as disaster rescue and prevention are also possible, as long as the nature of problem solving remains the same.

6.3. Limitations and suggestions

One major limitation of the proposed PPS is the requirement of the historic LLFs. Compilation of previous experience of problem solving is expensive. Moreover, the externalization of the tacit knowledge through Nonaka's knowledge creation spiral [25] is usually difficult. It may limit the compilation of useful historic LLFs. The LLW provided by CII may be employed to establish the LLF repository [13,27]. In the case study, the LLFs are compiled manually by the questioners who obtained solution from the domain experts. It is recommended to build the functionality of automatic LLF compilation with the established KMS [28]. Moreover, the project final reports, plans, proposals, and other knowledge documents contain tremendous engineering experiences that are valuable sources for solutions in solving future problems. Automatic systems should be developed to compile the explicit knowledge stored in the abovementioned documents into the LLFs.

A second limitation while implementing PPS in the case study was that the LLFs were not classified correctly. The current timesheet classification system is primarily for bookkeeping of payroll, rather than for knowledge management or problem-solving. Misclassifications were found frequently in the case study. Such misclassifications can lead to malfunction of APA module. A more accurate timesheet and document classification system should be established in order to improve the performance of APA.

Another limitation of PPS found from the case study was the keyword database. The present research adopted the Chinese Knowledge Information Processing (CKIP) database provided by the Institute of Information of the Academia Sinica [22]. The CKIP database provides only commonly used Chinese keywords rather than the construction-specific keyword database. It is recommended to the A/E consulting firm who implements PPS to establish their own special purpose keyword database.

The system testing of the present research also faces difficulty since the domain experts were extremely busy. It was almost impossible to ask them to perform testing experiments for PPS. It was found that some

domain experts provided invalid questions for the Similar Set while conducting the testing experiment of the research. Those data may mislead the testing results and should be excluded. Some standard testing data for the domain problem of construction should be established for verification of the proposed system.

Finally, the verification was only conducted for the case A/E firm. The applicability of the proposed PPS to other construction organization deserves further studies. Since CECI Consultants Inc. is the top ranking local A/E consulting firm, the professional backgrounds of the staffs cover almost all possible technical domains related to engineering consultants. There is rare situation that when a posed problem cannot be solved by the in-house staff. However, if this situation does happen, an external Subject Matter Expert (SME) is hired by project base to solve the problem.

7. Conclusions and future works

This paper presents the development of a new problem-solving method, Proactive Problem-Solver (PPS), for the construction industry. The proposed PPS differentiates itself from the traditional reactive problem-solver approaches by providing a proactive mode of problem solving. Such proactive problem solving is realized with the integration of the following components: a Multi-dimensional Knowledge Ontology (MKO) representation of the historic Knowledge Map (KMap), a Knowledge Capacity Matrix (KCM) scheme for characterization of the domain Expert Map (EMap), an Automatic Problem Answering (APA) module for retrieval of historical Lesson-Learned Files (LLFs), an Automatic Problem Dispatching (APD) module for diverting unsolved problems to the most appropriate domain experts of the firm, and a repository for storing the historic Lesson-Learned Files. From the case study results of a local leading A/E consulting firm in Taiwan, it is found that both timeliness and cost-saving performances are significantly improved. The *Recall* ratio of correct answers is 89.6% for the Similar Set (similar but articulated differently questions) as the number of retrieved LLFs is set to be 10. The timeliness efficiency of problem solving is improved by 89.5% for the Similar Set. With the proposed PPS, the problems encountered by construction engineers and managers in their daily operations and works can be solved more efficiently and effectively. It is concluded that the proposed PPS has a great potential for improvement of construction problem solving.

Although PPS shows promising potentials, the case study also found some limitations of the present version of PPS including: (1) the keyword database is not construction-specific, which causes the incorrect retrievals of historic LLFs; (2) the scope of historic LLFs is limited to a special CoP (SOS) of the case A/E consulting firm, it should be expanded to include other CoPs; (3) only the historic LLFs is employed for problem solving, other explicit knowledge documents (e.g., project final reports, plans, proposals, etc) should be included to build the historic LLFs; (4) misclassifications of knowledge documents and timesheets are commonly found, a more accurate technical classification system should be implemented.

Future extensions of the proposed PPS are also identified such as integrating with the ubiquitous technology to enhance the real-time problem solving, application of PPS to proposal preparation, application of similar methodology to problem-solving in other fields (e.g., legal affairs, disaster rescue, etc.) Ambitious researchers are encouraged to pursue in those directions.

Acknowledgements

References

[1] H. Li, P.E.D. Love, Developing a theory of construction problem solving, Construction Management and Economics 16 (6) (1998) 721–727.
[2] W.D. Yu, C.T. Lin, C.T. Yu, S.J. Liu, H.C. Luo, P.L. Chang, Integrating emergent problem-solving with construction knowledge management system, Proceedings of the CME 25 Conference, University of Reading, Reading, UK, July 16 18, 2007 (10 pp).
[3] J. Lave, E. Wenger, Situated Learning: Legitimate Peripheral Participation, Cambridge University Press, New York, 1991.
[4] W.D. Yu, P.L. Chang, H.H. Yao, S.J. Liu, KVAM: model for measuring knowledge management performance of engineering community of practice, Construction Management and Economics 27 (8) (2009) 733–747.
[5] S. Kim, E. Suh, H. Hwang, Building the knowledge map: an industrial case study, Journal Knowledge Management 7 (2) (2003) 34–45.
[6] M. Mezher, M.A. Abdul-Malak, I. Ghosn, M. Ajam, Knowledge management in mechanical and industrial engineering consulting: a case study, Journal Management in Engineering, ASCE 21 (3) (2005) 138–147.
[7] C.H. Caldas, L. Soibelman, J. Han, Automated classification of construction project documents, Journal of Computing in Civil Engineering 16 (4) (2002) 234–243.
[8] P.G. Ioannou, L.Y. Liu, D. Antonio Jr., ACTS: advanced construction technology system, Source Document 84, Construction Industry Institute, Austin, TX, U.S.A, 1992.
[9] J. Loss, AEPIC project: update, Journal of Performance of Constructed Facilities, ASCE 1 (1) (1987) 11–29.
[10] N. Kartam, I. Flood, Constructability feedback systems: issues and illustrative prototype, Journal of Performance of Constructed Facilities, ASCE 11 (4) (1997) 178–183.
[11] B. McCullouch, B. Patty, Hypermedia constructability system, Proceedings Computing in Civil Engineering, ASCE, New York, NY, USA, 1994, pp. 1397–1404, (n 20.
[12] N. Kartam, Knowledge-intensive database system for making effective use of construction lesson learned, Proceedings Computing in Civil Engineering, ASCE, New York, NY, USA, 1994, pp. 1139–1145, (n 2).
[13] Construction Industry Institute, Constructability Implementation Guide, Austin, TX, U.S.A, 1993 (Publication 34-1).
[14] Y. Mohamed, S. AbouRizk, Technical knowledge consolidation using theory of inventive problem solving, Vol. 131, 2005, pp. 993–1001.
[15] Altshuller, G. (2000). *The innovation algorithm: TRIZ, systematic innovation and technical creativity*, L. Shulyak and S Rodman, translators, Technical Innovation Center, MA, USA
[16] W.D. Yu, H.H. Yao, S.J. Liu, P.L. Chang, Knowledge value adding model for quantitative performance evaluation of the community of practice in a consulting firm, Proceedings of The Sixth International Conference on Engineering Computational Technology (ECT 2008), Session VIII — Decision Making in Engineering Management, 2008, (Sept. 2-5, Athens, Greece, 18 pp).
[17] D.L. Waltz, An English language question answering system for a large relational database, Comm, ACM 21 (7) (1978) 526–539.
[18] J.C.R. Tseng, G.J. Hwang, Development of an automatic customer service system on the internet, Electronic Commerce Research and Applications No. 6 (2007) 18–29.
[19] J.W. Wu, Y.H. Tsou, C.K. Chiou, J.C.R. Tseng, Development of a Ubiquitous Virtual Tutoring Assistant System, Proceedings of the 2007 World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA), Vancouver, Canada, June 25-June 29, 2007, (8 pp).
[20] G. Salton, A. Wong, C.S. Yang, A vector space model for automatic indexing, Communications of the ACM 18 (11) (1975) 613–620.
[21] G. Salton, M. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, NY, USA, 1983.
[22] W.Y. Ma, K.J. Chen, Introduction to CKIP Chinese Word Segmentation System for the First International Chinese Word Segmentation Bakeoff, Proceedings of the Second SIGHAN Workshop on Chinese Language Processing, Sapporo, Japan, July 11~13, 2003, pp. 168–171.
[23] W.D. Yu, P.L. Chang, S.J. Liu, Quantifying benefits of knowledge management system: a case study of an engineering consulting firm, Proceedings of International Symposium on Automation and Robotics in Construction 2006 (ISARC 2006), Session A4—Planning and Management, 1, Oct. 3~5, 2006, (Tokyo, Japan, 6 pp).
[24] C.W. Cleverdon, The Cranfield tests on index language devices, Aslib Proceedings No. 19 (1967) 173–192.
[25] I. Nonaka, A dynamic theory of organizational knowledge creation, Organization Science 5 (1) (1994) 14–37.
[26] F. Johansson, The Medici Effect: Breakthrough Insights at the Intersection of Ideas, Concepts, and Cultures, Harvard Business School Press, Cambridge, MA, USA, 2002.
[27] D. Fisher, S. Deshpande, J. Livingston, Modelling the Lessons Learned Process, Research Report 123–11, Construction Industry Institute, Austin, TX, USA, 1998.
[28] W.D. Yu, Y.T. Wei, S.J. Liu, P.L. Chang, Preliminary study on the automatic lessons-learned file generator, Proceedings of International Symposium on Automation and Robotics in Construction 2008, Vilnius, Lithuania, June 27–29, 2008, (ISARC 2008, 8 pp).