

IDENTIFYING CONSTRUCTION PROBLEM-SOLVING PATTERNS OF LESSONS LEARNED WITH TEXT MINING METHOD

Pei-lun Chang^{1*}, Wen-der Yu², and Shun-min Lee³

¹ *CECI Engineering Consultants, Inc., Taiwan, Taipei/ Chung Hua University, Hsinchu, Taiwan*

² *Chung Hua University, Hsinchu, Taiwan*

³ *CECI Engineering Consultants, Inc., Taiwan, Taipei, Taiwan*

* *Corresponding author (peilun@ceci.com.tw)*

ABSTRACT: Many engineering consulting firms have adopted knowledge management systems (KMSs) as a tool to record the knowledge and experiences of their engineers and staffs. It was found from previous research that similar construction problems were encountered and resolved by the same firm. The most commonly adopted form of such kind of knowledge and experiences is historical lesson-learned files (LLFs). It is very beneficial to identify the underlying patterns of historical LLFs, so that the retrievals and reuses of such knowledge and experiences can be more efficient and effective. This paper presents a text mining method that expedites the identification of construction problem-solving patterns from 908 historical LLFs recorded in the KMS of the case engineering consulting firm. Various text mining algorithms are tested to find out the appropriate ones that are most effective in the mining of problem-solving patterns. Domain experts are consulted to verify the results of text mining. The reapplication of the identified problem-solving patterns is demonstrated to compare the proposed method with the existing method both in efficiency and effectiveness of construction problem solving. The results of the present research show that the text mining and data mining methods have the potential to provide a solution for identifying the patterns of construction problem solving. With such patterns, the construction engineers and managers are better equipped and supported as they are encountered with the numerous emergent problems in their daily works.

Keywords: *Problem Solving, Knowledge Management, Lessons-learned, Text Mining, Data Mining*

1. INTRODUCTION

Knowledge management (KM) plays a more and more important role in both accumulating the organizational intellectual assets and solving the engineering problems for construction industry. Systematical implementation of KM methods for a construction organization does not only reduce the impact of knowledge leaks due to leaves of staffs, but also add the value of the personnel of the organization by leveraging the usage of that knowledge through knowledge management systems (KMSs). One successful application of KMSs is the collection and application of the historical lessons-learned in engineering problem solving, where the questioner poses a question on the community of practice (CoP) of the KMS and participants respond with their answers. It was found that

the KMS provides a forum for the solution-knowers with different context to provide their knowledge for solving the target problem from different angles. Such a phenomenon is called the Medici Effect [1]. The solution is then compiled by the questioner to become a lesson learned file (LLF) and is stored in the knowledge base of the KMS [2]. It is found from previous LLFs that problem-solving activities were repeated by the engineers and managers in their daily works. However, application scenarios of the LLFs may differ from case to case and from one discipline to another. It is also found that the problem of a specific discipline may be solved by the LLF of the other. As a result, it is very beneficial to identify the underlying patterns of historical LLFs, so that the retrievals and reuses

of such knowledge and experiences can be more efficient and effective.

The current research aims at identifying the underlying problem-solving patterns of the historical LLFs. A novel approach called text mining is adopted to assist the recognition and regulation of the underlying patterns. Special algorithms of text mining are developed to extract critical terms of the descriptions of the problem and solution. The data mining methods such as clustering, association rule, and decision trees were adopted to generate the problem-solving patterns from the collected LLFs.

The rest of the paper is structured as follows: the essentiality of engineering problem solving is discussed in Section 2. It is followed by the lesson-learned method that is usually adopted by engineering organizations for collecting problem-solving knowledge and experiences. In Section 4, the emergent problem-solving system of the case A/E firm selected for the case study in this paper is introduced. Then, the collected historical lessons-learned are described in Section 5. In Section 6, the Association Rule generation methods are adopted to generate the underlying problem-solving rules of the collected LLFs. Finally, conclusions are drawn from case study results and recommendations are made for future research.

2. PROBLEM SOLVING

Engineering consultant is a knowledge-based service industry. The knowledge creation, storing, accumulation, retrieval, and reusing are critical to the competitiveness of the engineering consulting firm. Problem solving has played the central role of the many business operations of an engineering consulting firm. All activities (e.g, planning, design development, project management, construction supervision, etc.) of an engineering consulting firm are associated with problem solving [2]. As a result, improvement of the effectiveness and efficiency of the problem-solving activities will definitely improve the competitiveness of the firm.

Smith [3] and Li and Love[4] have addressed the essentiality of construction problem solving. Li and Love[4] found that construction problems pose several

characteristics that should be tackled in order to solve them quickly, correctly, and cost-effectively; these characteristics include: (1) ill-structure nature—thus experimental knowledge plays important roles; (2) inadequate vocabulary—thus communications between researchers and practitioners is important; (3) little generalization and conceptualization—first solution is usually adopted, no guarantee on optimal solution; (4) temporary multi-organization (TMO)—relevant organizations have to work together in order to reach a consensual solution for all parties; (5) uniqueness of problems—it is hard to accumulate experiential knowledge from construction practices; and (6) hardness in reaching the optimal solution—adequate measures are required to evaluate the performance of problem solving, including quality of resultant solutions and their benefits. In Li and Love's research, they found that the abovementioned characteristics are generally tackled in two areas of problem solving researches: the cognitive science and decision support system (DSS). No quantitative measures were provided on how well the two approaches have accomplished in real world implementations.

Yu et al. [5] conducted a research on application of KMS for solving emergent problems. They proposed a Knowledge Management integrated Problem-Solver (KMIPS) for emergent problem solving of an engineering consulting firm [5]. They found that averagely 50.88% time-saving, 63.53% man-hour decreasing, and 84.40% cost effectiveness were achieved with KMIPS.

Dave and Koskela[6] believed that the problem-solving activity itself is not so challenging while compared with the documentation and reuse of the lesson-learned. Some researchers emphasized on the revision and creation of new knowledge (or so-called "dynamic knowledge") in order to improve the effectiveness of problem solving [7]. Others addressed that the vagueness and uncertainty of the application scenario of a lesson-learned may affect the performance of the problem solver [8].

Among the many characteristics of a construction problem, emergency may be the most important and influential one. Since most construction problems are urgent, late solutions may be invalid or even useless. For example, the piping of

ground soil may result in a failure or even a disaster of a building foundation work. As a result, a SOS (a specialized emergent problem solving subsystem of the KMS) has been developed to meet the requirement of timeliness for problem solving [9]. Based on the SOS system, Yu et al. proposed a Proactive Problem Solver (PPS) [10] and Integrated Proactive Knowledge Management Model (IPKMM) [11] to improve the performance of such kinds of KMS enhanced problem solvers.

In spite of the great efficiency and effectiveness achieved, the essential problems (the vagueness and uncertainty of the application scenario of a lesson-learned) encountered with the abovementioned systems has hindered further applications of KMS-enhanced problem solvers. In order to conquer the aforementioned problem, the essentiality and nature of construction lesson-learned are analyzed in the next section.

3. LESSON LEARNING

As a highly knowledge-intensive industry, the construction organization and enterprise need a better way for knowledge management. The lesson learning has been adopted by many construction organizations to record and reuse the previous experiences of problem solving [12]. According to Rosina Weber [13], lesson learning has been adopted by many organizations including military, business, and government to acquire, store, diffuse, and share knowledge since 1980's. Stewart Graham and Thomas address that the recording of knowledge and experience plays an important role in improving the performance of a project management team from their case study on an Irish leading construction firm [12]. Applications of lesson learned files (LLF) can be found in many real world examples, such as the Hypermedia Constructability System (HCS) developed by Indiana Department of Transportation (INDOT) and Purdue University[14], the Constructability Lessons Learned Database (CLLD)[15] & Integrated Knowledge-Intensive System (IKIS)[16] developed by Kartam and Flood, and the Lessons-Learned Wizard (LLW) developed by the Construction Industry Institute (CII).

Fisher et al. [17] proposed a lessons-learned process (LLP) for human problem solver, see Fig. 1. In Fig. 1, the information is collected as “captured information” first; then, it is analyzed and validated to become “knowledge”; only when the knowledge is implemented in the real world case, it becomes a “lesson-learned”. The vagueness and uncertainty of the application scenario, as addressed by Chang and Chui [8], may be generated during the collection and implementation stages described in Fig. 1.

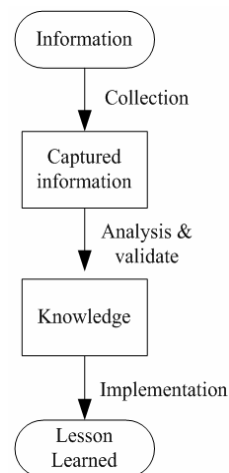


Fig. 1 Conceptual lesson-learned process [17]

4. SOS SYSTEM FOR PROBLEM SOVLING

A LLF system has been implemented by the case architect & engineering consulting firm (A/E firm), CECI Engineering Consultants, Inc.[18] in 2006. The LLF system adopted the LLW originally proposed by CII to record the problem-solving experiences recorded for the SOS system [9], as shown in Fig. 2. Once the emergent problem is solved with the SOS system, a LLF is compiled as shown in Fig. 3.

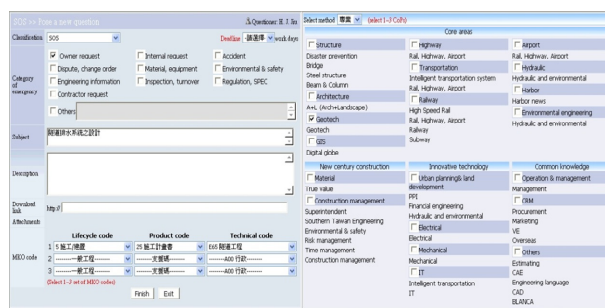


Fig. 2 The entry portal of SOS System

Name of form: SOS_Temp_01_P01		Lesson-Learned File	
1. Knowledge Classification (Library code: Digit 1, Product code: Digit 2-3, Technical code: Digit 4-6)			
Classification: 02007			
Keywords: 鋼筋			
2. Description of problem			
Subject: 鋼筋工程品質管理與檢驗			
Category of emergency:	Category of emergency: 品質管理與檢驗	Deadline:	Is problem solved?
Questioner: 43781	Time of Posting: 2008/10/17 11:11	Attachment:	Deleted link:
Description: 鋼筋工程品質管理與檢驗			
3. Description of the lesson learned			
Who help solve the problem: Department: SI, Name: 0505			
Date of problem solved: 2008/10/18 11:11			
Description of solution: 鋼筋工程品質管理			
Consideration grade: 3			
Attachment:			
4. Evaluation of the lesson learned			
Qualitative description of the benefit (optional):			
Benefit type:	Benefit description (optional):		
Benefit of problem-solving:	改善		
Cost benefit:	改善		
Time benefit:	改善		
Quality benefit:	改善		
Environmental and safety benefit:	改善		
Technology innovation or improvement:	改善		
Effect on regulations, SPEC, etc. (optional):			

Fig. 3 An example of Lesson-Learned File (LLF) obtained from SOS system

6. TEXT MINING OF LESSON LEARNED FILES

This section describes text mining process of the LLFs generated by the SOS system shown in the last section.

6.1 LLF Classification

There are totally 908 LLFs collected in the past six year via SOS. The 908 LLFs are categorized into 8 classes: (1) Specifications—256 cases; (2) Construction technology and management—244 cases; (3) Material testing—102 cases; (4) Contract—78 cases; (5) Data of vendors—61 cases; (6) Terminology and definitions—37 cases; (7) Computer and information technology—27 cases; and (8) Others. In this preliminary study, only “(2) Construction technology and management (244 cases)” and “(4) Contracts (78 cases)” are included for further study, since the other cases are requests for information rather than problem-solving cases. As a result, totally 322 LLFs are further analyzed, which quotes 35.46% of the total cases collected.

6.2 Analysis Procedure

The analysis procedure of proposed method is depicted in Fig. 4 and is comprised of the following steps: (1) Input LLFs; (2) LLF structure analysis; (3) Segmentation; (4) Key phrase extraction; (5) DM parameter setting; (6) Data mining; (7) Problem-solving pattern identification; (8) Pattern verification; and (9) Output problem-solving patterns.

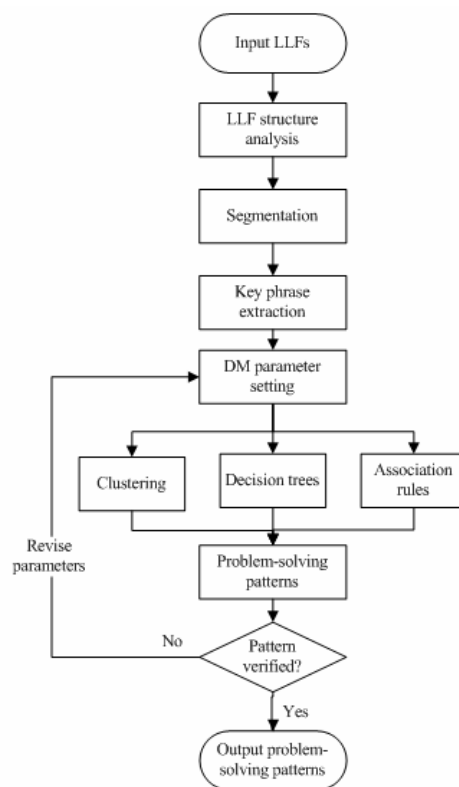


Fig. 4 Analysis procedure

6.3 Keyword Identification

In this research, the keyword databases from three different sources are collected including: (1) the CKIP keyword base—a corpus base storing five million most commonly used Chinese words and phrases provided by the Institute of Information Science, Academia Sinica of Taiwan [18]; (2) Construction professional keywords provided by the Public Construction Commission, Executive Yuan, Taiwan; and (3) the technical keywords provided by case A/E firm. The segmentation scheme adopted is “long-term first” algorithm. “Stop words” such as “thank you”, “please”, etc. are excluded from the keywords. Finally, totally 4,180 keywords are identified for “Problem descriptions” and 3,922 keywords are identified for “Solution descriptions” of the selected 322 LLFs.

6.4 Data Mining

The DM was performed with the Microsoft SQL Server® 2005. The general data DM procedure consists of the following 7 steps: (1) Adding a new analysis service

project; (2) Adding a new data source; (3) Linking the new data source; (4) Setting up data view of the new data source; (5) Selecting data table and data view; (6) Completing adding new data view; (7) Selecting a DM algorithm for data mining. In this paper, all nine DM algorithms provided by the Microsoft SQL Server[®] 2005 were tested with KMS performance data. Finally, three DM algorithms were adopted including Clustering, Decision Tree, and Association Rule. After DM experiments, it is found that only Association Rule is applicable. The other two methods did not generate meaningful results. Finally, 13 patterns are identified. Several examples of the identified patterns are shown for illustration in the following, for examples:

Pattern 1—If *ProblemType* is “Detailed Schematic Drawing” or “Price Related”, then *ProblemSolution* is “Follow construction price standards” or “Design Change” (Confidence=0.5, Importance=1.906).

Pattern 2—If *ProblemType* is “Solid Waste” or “Construction Time”, then *ProblemSolution* is “Schedule Extension” (Confidence=1, Importance=1.731).

Pattern 3—If *ProblemType* is “Public Arts” or “Change Order”, then *ProblemSolution* is “Negotiation” (Probability=1, Importance=1.634).

Pattern 4—If *ProblemType* is “Aerial Survey” or “Value Engineering”, then *ProblemSolution* is “Refer to Historical Data” (Confidence=1, Importance=1.634).

6.5 Pattern Verification

The identified problem-solving patterns described in the last section are presented to the engineers who have participated in developing the LLF to verify the appropriateness and the applicability of the patterns. The data mining has identified 108 patterns. However, only the rules with importance index > 1 were selected. As a result, totally 13 association rules were obtained. Finally, 13 patterns are verified and adopted as Engineering Problem-Solving Patterns of the case A/E firm.

7. CONCLUSION

This paper presents a text mining method for identification of the underlying problem-solving patterns of an engineering consulting firm. In this preliminary study, 322 historical problem-solving lesson-learned files are adopted as knowledge sources. Text mining methods were applied to identify the keywords of the technical domain. The commercial data mining software, Microsoft SQL Server[®], is adopted to find out meaningful patterns. Finally, the identified 13 patterns were verified by the domain experts who participated in solving the referred engineering problems. It is concluded that the proposed method provides a promising solution to construct an engineering problem-solving patterns that are universally applicable to solve generic engineering problems.

REFERENCES

- [1] Johansson, F., *The Medici Effect: Breakthrough Insights at the Intersection of Ideas, Concepts, and Cultures*, Harvard Business School Press, Cambridge, MA, USA, 2002.
- [2] Yu, W. D., Yang, J. B., Tseng, J. C. R., and Yu, C. T., “Model of Proactive Problem-Solving for Construction Knowledge Management,” *Proceedings of International Symposium on Automation and Robotics in Construction 2007 (ISARC 2007)*, Session 1-C—Management and Production Aspects (1), Sept. 19~21, 2007, Kochin, India, pp. 493~500, 2007a.
- [3] Smith, G.F., *Towards a Theory of Managerial Problem-Solving*. *Decision Support Systems*, 1992. 8(1): p. 29-40.
- [4] Li, H. and P.E.D. Love, *Developing a theory of construction problem solving*. *Construction Management and Economics*, 1998. 16(6): p. 721-727.
- [5] Yu, W. D., Lin, C. T., Yu, C. T., Liu, S. J., Luo, H. C., and Chang, P. L. “Integrating emergent problem-solving with construction knowledge management system,” *Proceedings of the CME 25 Conference*, July

- 16~18, 2007, University of Reading, Reading, UK, 10 pp., 2007.
- [6] Dave, B. and L. Koskela, Collaborative knowledge management-A construction case study. *Automation in Construction*, 2009. 18(7): p. 894-902.
- [7] Chinowsky, P., K. Molenaar, and A. Realph, Learning organizations in construction. *Journal of Management in Engineering*, ASCE, 2007. 23(1): p. 27-34.
- [8] Chang, A.S. and S.H. Chiu, Nature of engineering consulting projects. *Journal of Management in Engineering*, ASCE, 2005. 21(4): p. 179-188.
- [9] Chen, K. L., Chou, C. T., Liu, S. J., The Second Phase of KM Revolution in CECI—Implementation of Knowledge Growth Accounting, CECI Technological Records, p.186-199, 2006.
- [10] Yu, W. D., Yang, J. B., Tseng, J. C. R., Liu, S. J., and Wu, J. W. “Proactive Problem Solver for Construction,” *Automation in Construction*, Vol. 19, No. 6, p. 808-816, 2010.
- [11] Yu, W.D., et al., Enhancing Engineering Services with Integrated Proactive Knowledge Management Model, in the 14th International Research Symposium on Advancement of Construction Management and Real Estate (CRIOCM 2009). 2009, Advancement of Construction Management and Real Estate: Nanjing, China. p. 2049-2416.
- [12] Graham, B. and K. Thomas, Lessons Learned Practices: Case Study of a Leading Irish Construction Organization, in *Construction Management and Economics 25th Anniversary Conference*. 2007, University of Reading, UK.
- [13] Weber, R., D.W. Aha, and I. Becerra-Fernandez, Intelligent Lessons Learned Systems. *International Journal of Expert Systems*, 2001.
- [14] McCullouch, B., and Patty, B., Hypermedia constructability system, *Proceedings Computing in Civil Engineering (New York)*, ASCE, New York, NY, USA., n 2, p. 1397-1404, 1994.
- [15] Kartam, N., and Flood, I., Constructability feedback systems: Issues and illustrative prototype, *Journal of Performance of Constructed Facilities*, ASCE, Vol. 11, No. 4, p. 178-183, 1997.
- [16] Kartam, N., Knowledge-intensive database system for making effective use of construction lesson learned, *Proceedings Computing in Civil Engineering (New York)*, ASCE, New York, NY, USA., n 2, p. 1139-1145, 1994.
- [17] Fisher, D., Deshpande, S., and Livingston, J., Modelling the Lessons Learned Process, Research Report 123-11, Construction Industry Institute, Austin, TX, USA, 1998.
- [18] CECI, web site: <http://www.ceci.com.tw/>, visited 2011/03.
- [19] Institute of Information Science Web Site: <http://www.iis.sinica.edu.tw/>, Academia Sinica, visited 2011/3.