

A View from the Bottom: Project-Oriented Risk Mining Approach for Overseas Construction Projects

Lee, JeeHee¹ and Son, JeongWook^{2*} and Yi, June-Seong³

Abstract: Analysis of construction tender documents in overseas projects is a very important issue from a risk management point of view. Unfortunately, majority of construction firms are biased by winning contracts without in-depth analysis of tender documents. As a result, many contractors have incurred loss in overseas projects. Although a lot of risk analysis techniques have been introduced, most of them focus project's external unexpected risks such as country conditions and owner's financial standing. However, because those external risks are difficult to control and take preemptive action, we need to concentrate on project inherent risks. Based on this premise, this paper proposes a project-oriented risk mining approach which could detect and extract project risk factors automatically before they are materialized and assess them. This study presents a methodology regarding how to extract potential risks which exist in owner's project requirements and project tender documents using state of the art data analysis method such as text mining, data mining, and information visualization. The project-oriented risk mining approach is expected to effectively reflect project characteristics to the project risk management and could provide construction firms with valuable business intelligence.

Keywords: Overseas construction, Risk management, Project-oriented management, Text mining

I. INTRODUCTION

Construction tendering is one of the stages in a construction project that requires extensive information and documents exchange. Such tender documents often contain the information about a client's project plans so that a contractor can price it (Laryea 2011). Considering that construction project is a progressive work for materializing uncertain owner's requirements, tender documents are crucial and basic information which make project forward. It is important to both the owner and the contractor that the requirements for quality be clearly communicated in the RFP (tender document), so that the resultant proposals will be as responsive to the owner's needs as the cost, technical, and time constraints of the project allow (Gransberg & Molenaar 2004). However, close examination of large amounts of tender documents is always a challenging task due to the lack of tender period. Furthermore, tender documents are not always be clear. The discrepancies among the documents could cause severe problems in the construction phase. Therefore, analysis on tender documents should be managed from a risk management aspect.

Risk management is much more effective in the project early stage because initial risk management can take preemptive action. However, information uncertainty and the amount of available information is poor at a beginning step. Therefore, unclear and limited information at the beginning step should be detailed in order for maximizing initial risk management effects.

This study suggests a project-oriented risk mining approach which could extract potential project risks behind the tender documents especially in international projects. Since international projects are inherently exposed to

unpredictable and complicated risk scenarios (Han, S. H. et al. 2007), they might have many potential risk factors.

II. PROBLEM STATEMENT

A. An Overview of current research

Several different risk management researches have been performed in some aspects: definition of risk factors, building a risk response framework or system, and risk quantification for risk evaluation. The previous studies have weighted toward project external risk factors which is high-level risk, such as country risk or industrial risk. However, those external high-level risks are not easy to be managed, and does not reflect project specific risk as well. Moreover, most existing risk management and evaluation systems assess risks based on quantitative numerical value. However, much of the information related to risk analysis is not numerical. Rather, this information is expressed as words or sentences in a natural language (Kangari & Riggs 1989). That is the reason current risk management practice does not sufficiently deal with project specific risks in the real construction world. Besides, since majority of risk quantification studies relies on survey based statistical analysis which deals with generalized risk, the result from the analysis is difficult to reflect each project's characteristic.

Therefore, this study proposes project-oriented risk management approach which concentrate on project inherent risk factors (low-level risks) and documents based analysis using text analysis technology. Figure I shows the differences between current risk management studies and this study in terms of the level of risk factors (high-level risk factor, low-level risk factor) and risk management

¹ Graduate student, 52 ewhayeodae-gil, Seodaemun-gu, Seoul, 120-750, Korea, zehele@hotmail.com

² Assistant professor, 52 ewhayeodae-gil, Seodaemun-gu, Seoul, 120-750, Korea, jwson@ewha.ac.kr (*Corresponding Author)

³ Associate professor, 52 ewhayeodae-gil, Seodaemun-gu, Seoul, 120-750, Korea, jsyi@ewha.ac.kr

method (generalized risk management, project oriented risk management).

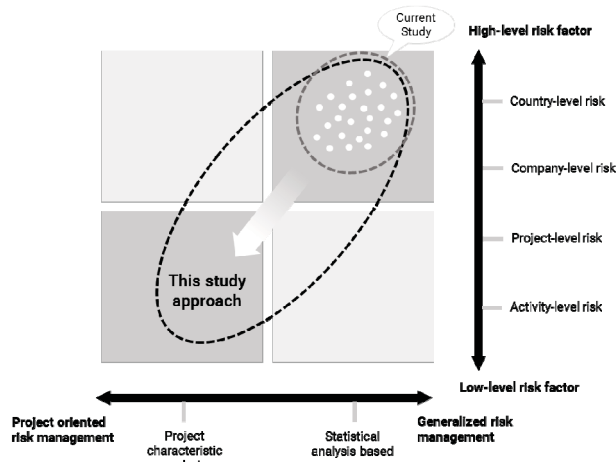


FIGURE I

DIFFERENCES BETWEEN CURRENT RISK MANAGEMENT STUDIES AND THIS STUDY

B. Research Questions

From the previous literature review, not only project external risk (high-level risk) factors but also project-oriented risk factors should be managed in the early stage of construction project. Because construction tender documents contain what the owner wants to build it, it could be a base information which make project forward. Based on this premise, the paper suggests a methodology regarding how to extract project-oriented risk factors from the tender documents and how to assess various types of risk factors. In order to proceed the study the following questions should be addressed:

- (1) What type of risks should/could be solved?
- (2) What kinds of project-oriented risks are existed in tender documents?
- (3) How to identify project-oriented risks?
- (4) How to extract/mine text-based risk factors?
- (5) What are the proper evaluation method for various types of risks?

III. RESEARCH METHODOLOGY

A. Project-Oriented Risk

Project-oriented risk could be defined as project specific risk factor which exist project inside. Project-oriented risk is related to the quality of tender document's information in drawings, specifications and bills of quantities included missing information, wrong information, insufficient detail and impracticable design, etc. Poor quality tender documents can lead to inaccurate estimates, higher margin in bids, claims and disputes (Laryea 2011).

There might be several ways to draw project-oriented risks in the overseas construction projects. The first way is reviewing preciously performed overseas construction projects. From the failure cases which experienced losses because of a poor analysis of tender documents, we could

understand what issues written on tender documents influenced the project's success and fail. Those issues might be the type of project-oriented risk factors. Another way to draw project-oriented risks is investigation into construction claim/disputes related to the tender document problems. Since claim and disputes are caused by project various risks, we could figure out project-oriented risk factors from claim and dispute cases.

Figure II shows risk identification and risk mining process. After various project-oriented risk factors are derived, risk vocabulary dictionary could be defined in order to identify project-oriented risk types. The defined risk vocabulary dictionary can make text-based documents possibly examined by risk mining methodology.

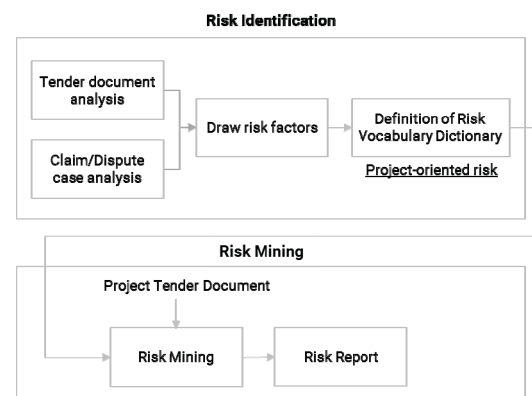


FIGURE II

RISK IDENTIFICATION AND RISK MINING

B. Risk Mining Process

This study defined the concept of risk mining as a progressive methodology which detect unstructured text data, and categorize risk types, and assess them. Implementing the risk mining should go through the following 3 steps: 1) text mining, 2) risk classification and 3) risk assessment.

Step 1: Text mining

Text mining is the first step of risk mining process which look for hidden information existed in tender documents as a text data form. Because the examination of tender documents has been conducted by some of experts so far, the quality of review was subject to be varied. Moreover, the analysis on tender documents has not been perfectly carried out due to the short tendering period and large amounts of documents. Thus, text mining could be a helpful method to analysis risks on documents and discover meaningful information. In order to mining text-based risk factors, definition of project-oriented risk factors should be performed in advance. Since text data is unstructured data, some of text-processing methods which transfer unstructured data into structured data should be conducted.

Step 2: Risk classification

Second step is risk classification which classify previously mined risks according to the risk type. In this step risk type is defined using prior performed project-

oriented risk factors. The risk type could be defined by risk predictability, risk controllability and level of risk, etc.

Step 3: Risk assessment

Final step is risk assessment which evaluate classified risk types in accordance with appropriate risk assessment method by risk type. Since the assessment method could be different by risk type which could be unstructured data or structured data, various data analysis method would be applied. In this step, the assessed risk type could be not only project-inside but also project-outside risk type. In other words, the initial analysis object is focused on project-oriented risks, which exist in tender documents, but the evaluated risk type from the risk mining process covers multi-level risk factors.

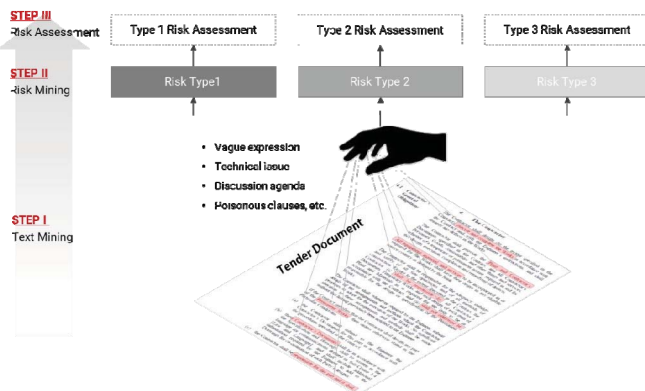


FIGURE III
RISK MINING PROCESS

C. Multilevel project-oriented risk mining methodology

According to the risk mining process above the multilevel project-oriented risk mining methodology could be illustrated as Figure IV.

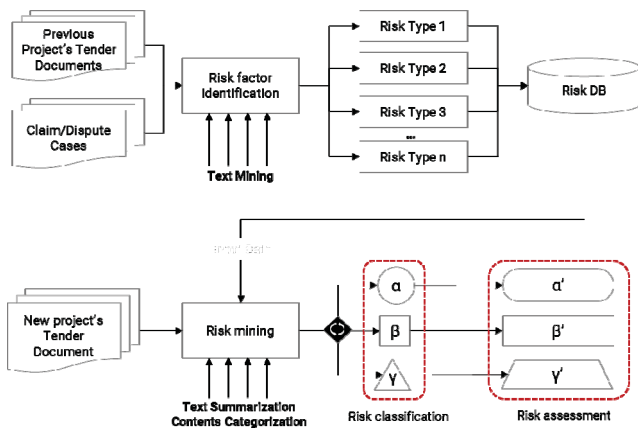


FIGURE IV
PROJECT-ORIENTED RISK MINING METHODOLOGY

Previously identified risk factor and types are stored in project-oriented risk DB, and the DB is used when the risk mining is in the process. In other words, the data stored in the DB is applied to the risk mining process as a mining subject. Thus, if newly incoming tender documents have similar words or phrases compared to the risk DB (risk

vocabulary dictionary), the words or phrases are detected, classified and assessed them by risk mining process.

IV. ILLUSTRATIVE EXAMPLE: TEXT MINING OF RISK FACTORS

To illustrate the proposed approach, analysis on construction law cases using text mining tool is performed on a trial basis to draw project-oriented risk factors which are related to construction claims and disputes. For the text mining analysis a total of 73 international construction law cases were collected, and NVivo 10 which is specialized for qualitative research was used for technical text processing.

A. Frequency Analysis

In order to derive risk-related words from the construction law cases, frequency analysis was conducted as a part of text mining. Since some words occurred most frequently, such as 'the', 'a', 'contract' and 'owner', etc., are not meaningful itself, these words were excluded on this frequency analysis. Table I shows the top 30 frequent words appeared in the law cases. There are some words related to cause of dispute such as 'payment', 'subcontract', 'breach', 'site', 'delay' and 'language'. These words also means risk-related words which should be examined during the tendering process. Although the frequency rate is not much high, however, important risk-related words could be placed low ranking. In other words, high frequency rate dose not stand for the importance. Therefore, it is necessary to search word relationship which shows the way how word appears with other words in the text-based documents. The following section (word tree analysis) gives some idea for understanding word relationship.

TABLE I
FREQUENCY ANALYSIS RESULT

No.	Word	Frequency	Weighted Percentage (%)
1	Court	88	2.4
2	Ruled	52	1.4
3	Appeal	44	1.2
4	Payment	34	0.9
5	Awarded	32	0.9
6	Conditions	30	0.8
7	Subcontract	30	0.8
8	Sued	29	0.8
9	Claim	28	0.7
10	Arbitration	26	0.7
11	Failed	24	0.7
12	Final	23	0.6
13	Agreement	22	0.6
14	Low	21	0.6
15	Refused	21	0.6
16	Argued	20	0.5
17	Breach	20	0.5
18	Entitled	20	0.5
19	Completion	19	0.5
20	Site	19	0.5
21	Stated	19	0.5
22	Delay	18	0.5
23	Price	18	0.5
24	Prime	18	0.5

25	Required	18	0.5
26	Supreme	18	0.5
27	Recover	17	0.5
28	Clause	16	0.4
29	Pay	16	0.4
30	Language	15	0.4

B. Word Tree Analysis

Word tree analysis displays the results as a tree with branches representing the various contexts in which the word or phrase occurs. Some words which are related to tendering terms (clause, agreement, failed etc.) were selected to make word tree in order to know the relations among words in a sentence. Figure V shows the word tree of 'clause' and 'failed'. According to the 'clause' word tree, 'no damage for delay' clause and 'order of precedence' clause are related to construction dispute, that is, these terms should be examined in advance for preventing risk arising. The word tree of 'failed' represents what types of dispute factors are existed in construction law cases. It could be possible to understand that 'giving timely notice' is also connected to the construction risk issue and should be managed at the tendering process. Using the word tree analysis more practical risk words and phrases could be derived, and based on the risk-related terms risk vocabulary dictionary could be defined.

clear, identifying owner's design intent written in tender documents is crucial for proceeding construction project. However, current risk management practice in the early stages does not sufficiently deal with project risks. In this paper, multilevel project-oriented risk mining methodology is presented based on investigation of current state of risk management. Moreover, to illustrate the proposed approach, mining risk factors from the construction law cases was performed. Because the illustrative example was only a part of drawing project-oriented risk factors, the text mining results are not enough to make project-oriented risk vocabulary dictionary and risk DB. However, if the study of previously performed project's tender documents will be conducted later, more practical risk factors could be derived. Using the risk factors risk mining process could be processing with computer-aided skills in the future study.

REFERENCES

- [1] Laryea, S., Quality of tender documents: case studies from the UK. *Construction Management and Economics*, 2011. 29(3): p. 275-286.
- [2] Gransberg, D. and K. Molenaar, Analysis of Owner's Design and Construction Quality Management Approaches in Design/Build Projects. *Journal of Management in Engineering*, 2004. 20(4): p. 162-169.
- [3] Han, S.H., D.Y. Kim, and H. Kim, Predicting profit performance for selecting candidate international construction projects. *Journal of construction engineering and management*, 2007. 133(6): p. 425-436.
- [4] Kangari, R. and L.S. Riggs, Construction risk assessment by linguistics. *Engineering*

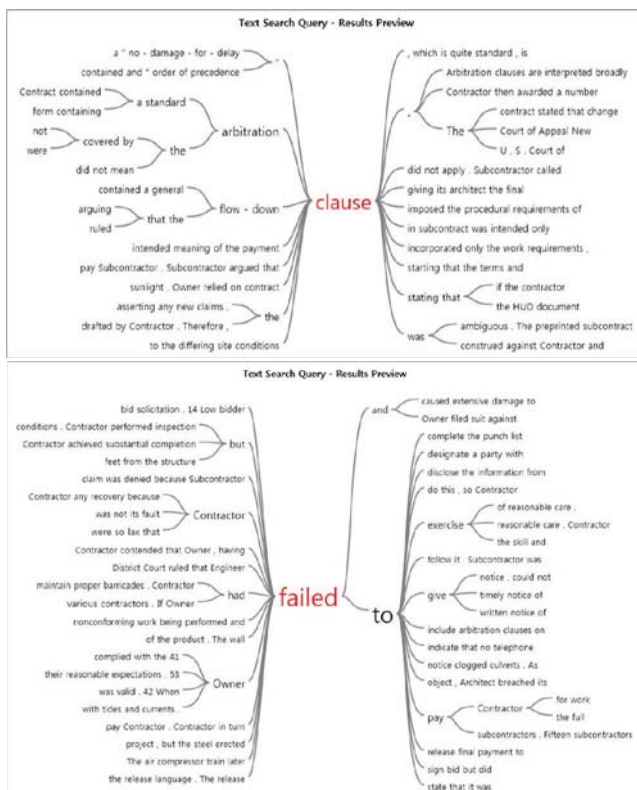


FIGURE V
 WORD TREE ANALYSIS

V. CONCLUSION AND FUTURE STUDY

The purpose of this paper is to propose a new approach for managing text-based project-oriented risks. Since available information in the early stages is not enough and