

Special Session 4

- Emerging Big Data Paradigms in the AEC Industry

건설 산업에서의 빅데이터 기술 활용 방안과 전망 -



구본상 서울과학기술대학교 건설시스템공학과 교수

서론

정보 시대의 '원유'라고 비유되는 빅데이터(Big Data)는 대용량 데이터 속에 숨겨진 패턴과 상호 연관성을 파악해 새로운 지식과 인사이트를 발굴할 수 있는 기회를 제공해 주고 있다. 컴퓨터 하드웨어 가격의 하락과 하둡(Hadoop)과 같은 병렬처리 기술의 발달로 인해, 빅데이터 기술은 더 이상 슈퍼 컴퓨터를 소유한 소수의 전유물이 아니라, 일반 회사와 개인까지도 자체 보유하고 있는 데이터의 분석을 가능케 한다.

빅데이터는 제조, 도소매, 금융, 의료, 엔터테인먼트 등 거의 모든 산업 부문에서 전방위적으로 그 활용도가 높아지고 있으며, 빅데이터 서비스(Big Data As a Service, BDaaS)도 하나의 신규 산업으로 발전하고 있다.

최근에는 건설산업 분야에서도 그 중요성이 부각되고 있으며, 해외에서는 건물 성능 및 에너지 분석, 인프라 자산 관리, 도시 개발 및 운영 등에 적용이 증대되고 있다. 또한 건설 프로젝트의 리스크 예측 및 클레임 방지, 현장의 안전사고 예측 등 다방면에서 연구 및 시도가 이어지고 있다.

본 세션에서는 국내에서 수행되고 있는 건설 분야 관련 빅데이터 연구를 소개하고 향후 전망을 살펴보았으며, 이를 통해 건설산업에 새로운 기회와 가치를 제시하고자 하였다.

다음은 각 발표자의 내용을 요약하였다.

- **텍스트마이닝 기법을 활용한 건설 문서로부터의 지식 추출 Creating Knowledge from Construction Documents Using Text Mining**
 - Seokho CHI (Seoul National University, Korea, Republic of)

건설산업에서 생성되는 중요하고 유용한 지식정보는 다양한 종류의 문서에 기록되고 있으며 그 양 또한 방대하다. 과거에 수행된 건설 프로젝트에서 발생한 문서에 포함된 지식정보는 타 유사 건설 프로젝트의 리스크 또는 전반적인 프로세스 관리 등에 활용될 수 있어 의사결정 및 사업전략수립을 위해 매우 중요한 역할을 한다. 따라서 건설 프로젝트 수행의 효율성을 향상시키기 위하여 건설 문서에 포함되어 있는 지식정보를 실무에서 사용 가능한 형태로 추출해내는 작업은 반드시 필요하다. 본 연구는 텍스트마이닝 기법을 이용하여 자동적으로 건설 문서로부터 지식을 추출해내는 시스템을 제안한다. 본 연구에서 제안하는 시스템은 건설 문서 전처리, 용어의 가치치 계산, 그리고 시각화의 세 주요 단계로 구성되었으며, 예비연구(pilot study)로 시스템 원형(原型)(prototype)이 개발되었다. 본 연구에서 개발된 시스템 원형은 텍스트마이닝 기법과 시각화에 기반한 지식 발굴 시스템의 유용성 및 가능성을 입증하였다. 자동화된 건설 문서 내용의 시각화는 기존에 존재하던 문서를 해석하기 편리한 형태로 가공하고, 방대한 양의 문서를 읽고 그 핵심 내용을 파악하는데 소모되는 시간과 에너지를 크게 줄여주었을 뿐만 아니라 건설 사업관리에 대한 통찰력을 제공하는 측면에서도 도움을 주었다.

- **해외 건설 사업 리스크 감식을 위한 텍스트 마이닝 분석 A Multilevel Project-Oriented Risk-Mining Framework for Overseas Construction Projects**
 - JeongWook SON (Ewha University, Korea, Republic of)
- 국내 건설기업의 해외 시장 진출이 확대되면서, 해외건설 프로젝트의 리스크 관리에 대한 중요성이 증대되고 있다. 그러나 지금까지 많은 프로젝트 초기단계의 리스크 평가를 위한

선행연구의 대부분이 국가 리스크, 발주자 리스크, 재무적 리스크 등 건설기업에서 통제 불가능한 리스크에 초점을 맞추고 있어 프로젝트 특성을 반영한 리스크 평가가 이루어지지 못하고 있다.

이에 본 연구에서는 입찰문서 내에 잠재되어 있는 프로젝트 리스크 요인들을 탐지하고 추출하여, 리스크가 발현되기 전에 평가할 수 있는 프로젝트 기반의 리스크 마이닝 방법론(Project-Oriented Risk Mining)을 제안하고자 한다. 이를 위해 입찰문서 내에 잠재된 리스크 요인들을 추출하고 파악하기 위해 텍스트 마이닝, 데이터 마이닝 및 정보 시각화 방법 등을 사용하여 문서기반의 리스크 평가를 실시하고자 한다. 본 연구에서 제안하는 프로젝트 기반의 리스크 마이닝 관리 방법은 기존의 리스크 평가 방식에서 가지고 있던 프로젝트 특성 반영의 어려움에 대한 한계를 극복하고, 실무 적용성을 높일 수 있는 실용화된 연구가 될 수 있을 것으로 기대한다.

■ 도시 빅데이터 (Urban informatics)를 통한 21세기 도시화 문제 해결 Forecasting Housing Demand with Big Data

- Bonsang Koo (Seoul National University of Science and Technology, Korea, Republic of)

최근 UN 보고서에 의하면, 오늘날 전세계 인구의 54%가 도시에 거주하고 있으며, 2050년에는 이 수치가 66%로 증가할 것으로 전망하고 있다 (United Nations, Department of Economic and Social Affairs, Population Division (2015)). 이를 전세계 인구 증가 전망 수치와 연계해서 보면, 2050년까지 약 33억명이 추가로 도심지역에 거주할 것으로 해석되고 있다. 이와 같은 도시 인구의 급증은 밀집화(overcrowding), 도시 빈곤, 도시 공해 및 교통 체증의 증가, 등과 같은 여러 가지 문제를 수반할 것으로 예측된다.

최근 빅데이터의 대두에 따라 이런 도시화의 문제를 데이터 기반으로 접근하려는 새로운 시도가 생기고 있으며, 이를 Urban Informatics, 또는 Urban Science로 명명하고 있다. Urban Informatics는 도시에서 생성되고 도시 정부에서 개방하는 각종 데이터(센서, 모바일 데이터, 공공 개방 데이터(Open Data) 등)를 활용하여 이를 도시 개발, 정책, 관리할 때 생기는 문제를 해결하려고 하는 노력이다.

Urban informatics는 학계의 연구에 국한돼 있지 않고, 실제 주요도시의 행정부와 국제기구에서 그 활용성을 적극적으로 실천하고 있다. 더불어 새로운 스타업 회사들도 창의적인 아이디어를 통해 문제를 풀어나가려고 하는 시도를 하고 있다. 본고에

서는 상기 사례들을 소개하면서 특히 건축 및 인프라 부문에 적용된 사례들을 소개하고 향후 이 분야에 전망과 우리가 접근할 수 있는 방안을 제시하고자 하였다.

■ 해외사업 잠재 리스크 파악을 위한 IPRI 지수의 개발 Development of International Project Risk Index (IPRI)

- Wi Sung YOO(Construction Economy Research Institute of Korea, Korea, Republic of)

2000년대 중후반부터 국내 대형 건설기업들은 해외시장에서 경쟁력 있는 건설 상품을 발굴하고 시장을 확대하기 위해 부단히 노력하고 있다. 국내건설시장의 지속적인 성장 정체로 인해 기업들은 글로벌 경쟁력 강화와 함께 수주 확대와 수익성 제고를 위한 전략 수립에 투자를 집중하고 있다. 하지만 최근 건설기업들이 해외건설시장에서 손실을 반복적으로 경험하면서 이윤의 극대화와 손실의 현실화를 예방하기 위해 조기경보기능을 확보하기 위한 대안을 모색하고 있다. 본 연구에서 국내 건설기업들이 수행하는 해외사업의 잠재 리스크를 파악하여 효율적으로 대응할 수 있도록 지원하는 리스크지수(IPRI)를 소개한다. IPRI는 3,500개의 사업 진행률 데이터를 추출 및 정제하여 개별 사업의 잠재 리스크를 계량화한 후 통합하였기 때문에 기업과 산업 차원에서 선제적으로 대응할 수 있는 방향성을 제시한다. 또한 정부 차원의 해외사업 금융 지원 정책을 수립하는데 기초 자료로 활용될 것으로 기대된다.

결론

본 세션에서 보듯이, 국내에서는 리스크 사업이 큰 해외 건설 시장 분석 및 해외건설 정보에 빅데이터 분석을 한 사례와, 건축 및 인프라 개발/관리 및 수요 예측에 적용된 사례들이 주를 이루고 있다.

최근 건설 경기가 어려우며 해외에서의 저가수주로 인한 막대한 손실을 드러내고 있다. 이는 데이터에 기반하지 않고 제한된 경험과 주관적 직관에 의존한 의사결정 프로세스도 분명 일조하고 있다.

빅데이터는 데이터의 크기가 중요한 게 아니라 데이터가 제공하는 노하우나 결과를 기반으로 합리적 의사결정을 가능케 한다는 데서 가치를 찾을 수 있다. 이런 맥락에서 건설 분야에서도 빅데이터의 활용이 늘어 더욱 성숙하고 고도화된 산업으로 성장하기를 기대해 본다.

Special Session 4 – Emerging Big Data Paradigms in the AEC Industry

Bon – Sang Koo

Professor of Department of Civil Engineering in Seoul National University of Science and Technology



Introduction

Big Data, so called the ‘newoil’ of the Information Age, provides opportunities to discover new knowledge by finding new patterns and insights hidden in massive data sets. Big Data technology is no longer the exclusive property of few people possessing supercomputers; due to the decline in computer hardware prices and development of parallel processing technology like Hadoop, companies of all sizes and even individuals can now analyze data in their possession.

The utilization of Big Data is expanding in all industries, including manufacturing, wholesale and retail, finance, medical, and entertainment. Also, the provision and service of Big Data techniques, so called, ‘BDaaS’ (Big Data As a Service), is itself developing into a new industry.

More recently, the application of Big Data has also been expanding into the AEC industry. Cases are emerging in which predictive analytics have been used in building capacity and energy analysis, infrastructure asset management, in city development and operations management. In addition, new research is emerging areas such as prediction of construction project risks, prevention of claims and prediction of on-site safety accidents.

In this session, domestic research on construction

related to Big Data is introduced, and its future prospects are examined. Through the session, new opportunities and potential values of presented to the AEC industry were explored. The contents of each presentation are summarized below.

■ Creating Knowledge from Construction Documents Using Text Mining

– Seokho CHI (Seoul National University, Korea, Republic of)

Important and useful knowledge information, created from construction industry, is recorded in various types of documents and its amount is also massive. Knowledge information included in documents from past construction projects has a very important role in decision-making and establishment of business strategy due to its use in other similar construction project risks or general business management. Therefore, process of extracting knowledge information included in construction documents into a form usable on-site is absolutely necessary to improve efficiency of construction project performance. This research suggests the system that uses text mining method to automatically extract knowledge from construction documents. The system suggested by this research is composed into three main steps: construction document preprocessing, weight calculation of terms,

and visualization. In a pilot study, a system prototype was developed. The system prototype developed from this research proved the usefulness and the potential of knowledge discovery system based on text mining method and visualization.

The visualization of automated contents of construction document turned the existing document into a convenient form for interpreting, and it not only drastically reduced the time and energy required for reading massive amount of documents and comprehending the main contents, but also assisted in providing insight for construction industry management.

■ **A Multilevel Project-Oriented Risk-Mining Framework for Overseas Construction Projects**

– JeongWook SON (Ewha University, Korea, Republic of)

As foreign market entries of domestic construction companies expand, the importance of risk management of overseas construction projects is increasing. However, risk assessment reflecting project characteristics has not been conducted as most advanced researches for early stage risk assessment of project are focusing on risks that are beyond construction companies' control, such as national risk, contractor risk, and financial risk.

Therefore this research examined and extracted potential project risks inside bid documents to suggest Project-Oriented Risk Mining that can evaluate risk prior to its occurrence. For this purpose, document-oriented risk assessment is to be conducted by using text mining, data mining, and information visualization methods to extract and comprehend potential risk factors inside bid documents. Project-Oriented Risk Mining, suggested in this research, is expected to overcome limits of the existing risk assessment methods, due to their poor reflection of project characteristics, and to become a commercialized research increasing practical

applicability.

■ **Forecasting Housing Demand with Big Data**

– Bonsang Koo (Seoul National University of Science and Technology, Korea, Republic of)

According to the recent UN report, 54% of the world population are living in cities today and this figure is expected to increase to 66% in 2050 (United Nations, Department of Economic and Social Affairs, Population Division (2015)). If these statistics are interpreted with estimations of world population growth, approximately an additional 3.3 billion people will live in cities by 2050. The rapid increase of urban population will include various problems such as overcrowding, urban poverty, and the increase of city pollution and traffic congestion.

Following the recent increasing influence of big data, new studies for approaching problems of urbanization based on data are conducted. This is defined as Urban Informatics or Urban Science. Urban Informatics is an effort to solve problems occurring in city development, policy, and management by using various data (sensor, mobile data, open data, etc.) created in cities and made public by city governments.

Urban informatics is not restricted to academic research and its usability is actively practiced in administration of major cities and international organizations. Additionally, new startup companies are attempting to solve problems through creative ideas. In this paper, cases of urban informatics application, especially those in construction and infrastructure, are introduced, and future prospects and alternative approaches of urban informatics are suggested.

■ **Development of International Project Risk Index (IPRI)**

– Wi Sung YOO (Construction Economy Research Institute of Korea, Korea, Republic of)

From mid-to-late year 2000, large domestic construction companies have been making ceaseless efforts to discover competitive construction products at foreign markets and to expand the market. Due to the continuous stagnation in the growth of domestic construction market, companies are focusing on the investment in strategy establishment for strengthening global competitiveness, order expansion, and profit improvement. However, as construction companies repeatedly experienced losses in foreign construction market, they are seeking alternatives to secure an early-warning system for profit maximization and to prevent losses. This research introduces IPRI, which comprehends the potential risks of foreign businesses performed by domestic construction companies and supports domestic companies to effectively respond.

As IPRI combines data after quantifying potential risks of individual businesses by extracting and refining progress rate data of 35000 projects, it presents directions which allows preemptive responses in the aspect of company and industry. In addition, it is expected to be applied as basic data for establishing financial support policies of foreign businesses in government level.

Conclusion

This session introduced cases in which Big Data analysis was conducted on foreign construction market studies, construction information possessing for high risk businesses, and cases in which big data analysis was applied to the development, management, and demand forecasting of construction and infrastructure in the urban environment.

The Korean construction industry is in a severe recession, and construction companies have suffered huge losses due to low-price bidding in international markets. The losses are in part due to subjective

judgements and biased decision making which has been the standard business-as-usual way of doing business. The value of Big Data is not in the size of the data, but rather in the possibility of allowing reasonable decision-making based insights and construction inferred from the data. In this context, it is hoped Big Data is embraced to advance and empower the AEC industry.