Big Data Analysis based Practical Applications in Construction

Yongho Ko[1], **and Seungwoo Han**[1*]
[1]Department of Architectural Engineering, Inha University, Incheon, South Korea

Abstract – *Big data analytic concept has changed fundamental paradigm of data analysis and prediction. Construction industry which seems to be the most conservative industry due to difficulty in data collecting, analyzing, and predicting is about to develop the advanced further phase with application of big data analytic methods. This study presents two cases which a big data analytic concept was adopted. The first case is for construction performance assessment and prediction. Home sales index prediction is the next. Database establishment with combination of simulation techniques and statistical approaches shows the fundamental ways how to use construction raw data stored in project management systems. Home sales index prediction based on search queries also suggests the effective methodology which enables users to investigate relevant environments and to produce reliable prediction results.*

Keywords: *construction performance, home sales index, prediction, assessment*

1 Introduction

The construction industry has been identified as one of the most conservative industries. It has been due to difficulties with emerging technologies applications in construction fields including house markets, where all conditions around each project have been totally different and various even very similar projects are operated. These difficulties have mostly been caused from limitations on the acquisition of data that could function as valuable input information on analytic and predictable methods [1]. For overcoming these limitations in construction, many studies related to searching efficient methods for collecting and analyzing informal data have been conducted in various fields in academia and industries. However, the effective ways for controlling and analyzing a huge amount of raw data which are collected in formal, informal, or semi-formal types have not suggested yet due to a sort of technical issue. The big data analytic concept significant findings in various fields currently get an attention as one of solutions for resolving these problems. The purpose of this study is to illustrate the preliminary cases illustrating how to apply the big data analytic concept to the construction. In accordance with this research objective, this study focuses two areas: 1) establishment of database for assessment and predictions of construction performances using data collected from construction sites [1, 2]; 2) home sales index prediction based on search query analysis. Figure 1 shows the diagram of two targeted areas in this study.

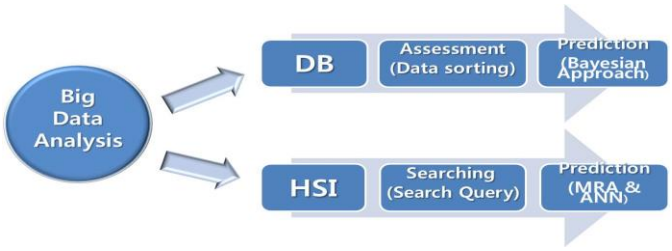

Figure 1. Two case studies for big data analytic concept applications

2 Construction Data

Difficulty in collecting precise raw data by consistent collection systems indicates why decision-making tools based on collected data have not been effective in the construction industry [1, 2]. This limitation in construction has brought other methodologies for data generations such as experimental designs based on statistic knowledge, and simulation techniques based on operational knowledge [1, 2]. Regarding predictions of construction performance, a multiple regression analysis presenting significant prediction results in various areas in academia and industry has been utilized as the representative methods based on the statistic knowledge. A fuzzy logic evaluating qualitative information to quantitative data and an artificial neural network also showing appropriate uses in prediction in industrial engineering have been mostly used based on artificial intelligences. However, it notes that these methods have not performed without collected raw data that reflects feasible cases under various conditions [1, 2].

3 Practical Applications

3.1 Database establishment for assessment and prediction of construction performance

Prediction of construction performance which is normally counted as fundamental criteria for making preliminary operational plans has been mostly studied since decision making tools were introduced in construction. Also precise predictions are critically demanding accurate assessment [1]. This case is for developing a database by using the simulation technique which has been recognized as an appropriate methodology for analyzing construction

performance but is limited in terms of its application to actual cases. The first phase of this database is to function proper assessment of construction performance data typically stored in project management information system as various types of data in formal, informal, and semi-formal formats. The second phase is then to predict performances using appropriate methodologies such as pattern recognitions and Bayesian approaches [2]. These two main phases in the database, assessment and prediction are conducted based on big data analytic methods. Figures 2 and 3 illustrate these two phases of output screen of the database [1, 2].

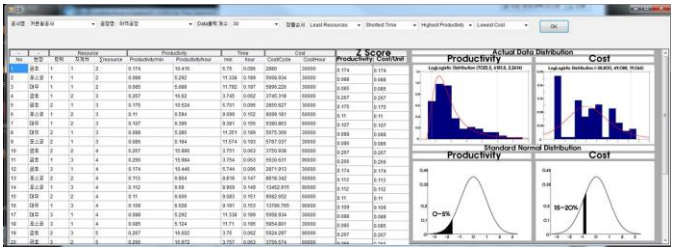

Figure 2. Output of the database for performance assessment [1]

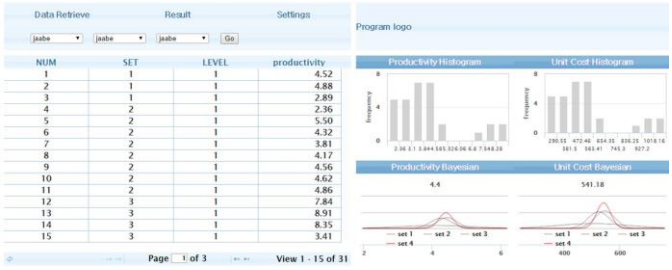

Figure 3. Output of the database for performance prediction [2]

3.2 Home sales index prediction

Home sales index has been identified to be one of key factors for establishment of long-term economic plans in government and of fundamental elements for preliminary design plans of construction projects. Subjective decision from empirical analysis and just experience of professionals were generally used in prediction of home sales index. This was derived from the difficulty and practical limitation posed by collection and analysis of a huge amount of relevant data. The case of home sales index prediction in this study utilizes the similar methods based on search queries in internet basis which was initiated by Google for detecting influenza epidemics [3]. This case shows a new home sales index prediction method based on a big data analytic concept implemented with statistical approaches such as cluster and principal component analysis. Multiple regression analysis and artificial neural network are applied for enhancing prediction results. Figure 4 illustrates the prediction results produced by these two methods.

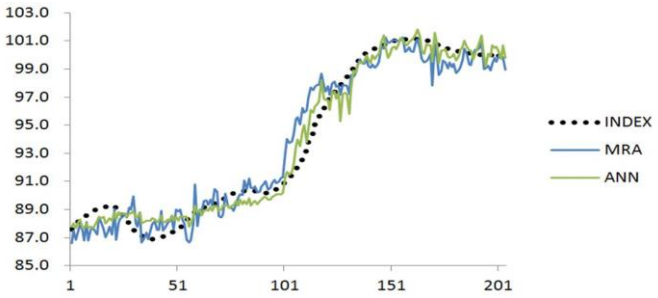

Figure 4. Prediction results

4 Conclusions

In construction, continuous efforts of development and practical application of appropriate decision making tools which were validated in other fields have been pursued. Several studies presented that big data analytic concept would be feasible for practical applications. This study suggested the practical cases for solving problems identified in the construction area especially practical applications. The cases are 1) construction performance assessment and prediction using raw data from project management information systems, and 2) home sales index prediction using search queries. This study presented the high possibility on a big data analytic concept of practical applications in construction with provision of significant findings and results. The methodologies and cases can be beneficial to other researchers interested in appropriate analytic and predictive methods in construction. It would be great initiations of the construction industry which make it more competitive against other manufacturing industries.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF), funded by the Korean government (NRF-2012R1A2A2A01046193)

5 References

[1] Seungwoo Han, Taehee Lee, and Yongho Ko. "Implementation of Construction Performance Database Prototype for Curtain Wall Operation in High-Rise Building Construction", Journal of Asian Architecture and Building Engineering, 13, 1, 149-156, Jan 2014

[2] Yongho Ko, and Seungwoo Han, "Development of Construction Performance Monitoring Methodology using the Bayesian Probabilistic Approach", Journal of Asian Architecture and Building Engineering, 143, 1, 73-80, Jan 2015

[3] Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant, "Detecting Influenza Epidemics using Search Engine Query Data, Nature, 457, 7232, 1012-1014, Feb 2009