Measuring the Underlying Structure of Classic Topics in Construction Engineering and Management: A SNA-based Analysis

Zebin ZHAO[1], Hengqin WU[2], Xiaolong XUE[3*], Geoffrey Qiping SHEN[4] , Boyu ZHANG[5] and Xin LI[6]

[1]Associated professor, Zebin ZHAO, School of Management, Harbin Institute of Technology, Harbin, China (corresponding author). Email: zhaozebin@hit.edu.cn.

[2]Ph.D. Candidate, School of Management, Harbin Institute of Technology, Harbin, China, 150001; Email: hengqin.wu@connect.polyu.hk

[3]Corresponding author, Professor, Xiaolong XUE, School of Management, Guangzhou University, Guangzhou, China; School of Management, Harbin Institute of Technology, Harbin, China; Email: xlxue@hit.edu.cn

[4]Chair Professor, Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong, China; Email: geoffrey.shen@polyu.edu.hk

[5]Ph.D. Candidate, Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong, China; Email: boyu.zhang@connect.polyu.hk

[6] Longgang Branch of Daqing Oilfield Information Technology Company, DaQing, China, 163453; Email: lix02@cnpc.com.cn

ABSTRACT

As an academic field, Construction Engineering and Management (CEM) emerged and developed into a mutual academic field in the past several decades. CEM is characterized by the high level of mixed knowledge from other disciplines (i.e. physics and computer science) to facilitate the application of engineering skills and models to execute the effective organizations in a complex environment. Therefore, understanding the latent topics in CEM is an important task. This study explores the classic topics of CEM and maps their structure by using SNA techniques from a large amount of data, providing relatively unbiased and convinced results.

INTRODUCTION

The landscape on the earth has been extremely changed by the construction industry, forming the built environment in which people could live with more convenience (Floricel et al., 2016). That build environment, gradually replacing the natural environment, was dominated by the construction and engineering processes that delivered physical construction products (such as buildings, roads, tunnels, bridges, airports, railroads, facilities, dams, and utilities) that occupied people’s living space. However, such a build environment was criticized by its poor performance (i.e. cost overrun, delays, benefit shortfalls and so on) (Flyvbjerg, 2014) and the negative impacts on the natural environment. Regardless of the concerns of project performance and sustainability, the impacts on the society (i.e. employment and other social responsibility) and economy (i.e. investment and benefits from the externalities of the infrastructure and other public facilities) should not be neglect. Therefore, construction engineering and management (CEM) has gained an increasing attention as an academic field, providing advanced knowledge for managers to guide their piloting of the projects, not only maximizing the value and reducing the risks, but also meet the requirement of preserving the natural environment and social responsibilities (Flyvbjerg, 2014; Laurance et al., 2014).

Recent development of mega-projects over the world involving huge complexities and uncertainties might lead to the requirement of advanced knowledge of CEM. One of the tasks, is to perceive the major themes of CEM and the associations between them. Indeed, understanding the theme structure in a given domain is a hard work, especially for non-experts, because the connections between topics are too implicit to trace and it would cost a large amount of time and energy to understand the logics underlying the syntax and linguistics in the context whey reviewing the archived documents (i.e. books and academic papers) (Foray and Lundvall, 1998; Silva et al., 2016). Such a dilemma highlights the need of big data-based approaches, which provides a novel research manner in which the latent knowledge could be quantitatively measured with visible and convinced findings.

This paper performs social network analysis (SNA) and data mining techniques, exploring the major topics in the field of CEM, and more important, measuring the connections between the topics. In specific, we extracted a large amount of data related to CEM, and identified the keywords, their counts and co-occurrence relations (Zhu et al., 2015). Price principle (Price, 1976) was used to filter the classic topics, and modularity optimization algorithms (Newman, 2004) made a partition of the theme network that was established based on the identified co-occurrence relations.

METHODOLOGY

(1) Data collection and processing

This study performs the topic structure based on analysis of the academic publications of CEM. Firstly, we retrieve the academic articles relevant to CEM, and extract all the keywords listed by the authors in those papers. To retrieve most relevant papers of CEM, 13 journals in the field of CEM were selected based on two main factors: (1) Journals in the ***Web of Science*** (WoS) are exclusively considered, because WoS is widely accepted by the scholars as an authoritative academic database. (Meho and Yang, 2007); (2) a journal has to be recognized as represented publications by the research community of CEM (Chan et al., 2004; Lin and Shen, 2007; Wing, 1997). Following these criteria, 13 journals were selected, as shown in Table 1. All the retrieved data have to be pre-processed, in which the keywords and their co-occurrence relations could be identified.

Table 1. 13 core journals of CEM

	Category	Journal	Publisher
1	Theory, method and application	Building Research and Information(BRI)	Taylor & Francis
2	Theory, method and application	IEEE Transactions on Engineering Management (IEEE M)	IEEE Xplore
3	Theory, method and application	Journal of Civil Engineering and Management(JCEM)	Taylor & Francis
4	Theory, method and application	Journal of Infrastructure Systems(JIS)	ASCE

5	Theory, method and application	Journal of Management in Engineering(JME)	ASCE
6	Theory, method and application	Journal of Construction Engineering and Management(CEM)	ASCE
7	Theory, method and application	Journal of Professional Issues in Engineering Education and Practice(PIEDP)	ASCE
8	Information & technology	Computer-Aided Civil and Infrastructure Engineering(CCIE)	John Wiley & Sons
9	Information & technology	Automation in Construction(AIC)	Elsevier
10	Information & technology	Advanced Engineering Informatics(AEI)	Elsevier
11	Information & technology	Journal of Computing in Civil Engineering(CCE)	ASCE
12	Project management	International Journal of Project Management(IJPM)	Elsevier
13	Project management	Project Management Journal(PM)	John Wiley & Sons

(2) Filtering out classic topics by Price Index

According to Price (1976), the number of classics (m) for a given scientific system follows the equation: $m = 0.749\sqrt{n_{maz}}$, where n_{maz} denotes all the items in the system (the number of the keywords in this case).

(3) SNA analysis

We utilized the identified classic keywords and their co-occurrence relations to build up the topic network of CEM. After that, a partition algorithm, modularity optimization, was used to divide the topics into sub-groups. Modularity optimization is an SNA-based algorithm to detect sub-groups (Newman, 2006). It is defined as (Newman, 2004):

$$Q=\frac{1}{2m}\sum_{i,j}\left[A_{ij}-\frac{k_i k_j}{2m}\right]\delta(c_i,c_j) \quad (5)$$

Generally speaking, modularity optimization performs a partition based on the links, categorizing the nodes into sub-group in which the internal nodes of every sub-group have density connections than external ones (Small, 1973).

RESULTS

According to the analysis procedure that was mentioned above, this study identified 122 classic topics (represented by keywords). A topic network is established to model the classic topics of CEM, in which the modularity optimization algorithm divided the network into subgroups, denoted as different colors (Figure 1). The size of each node represents the occurrence times of the corresponding topic.

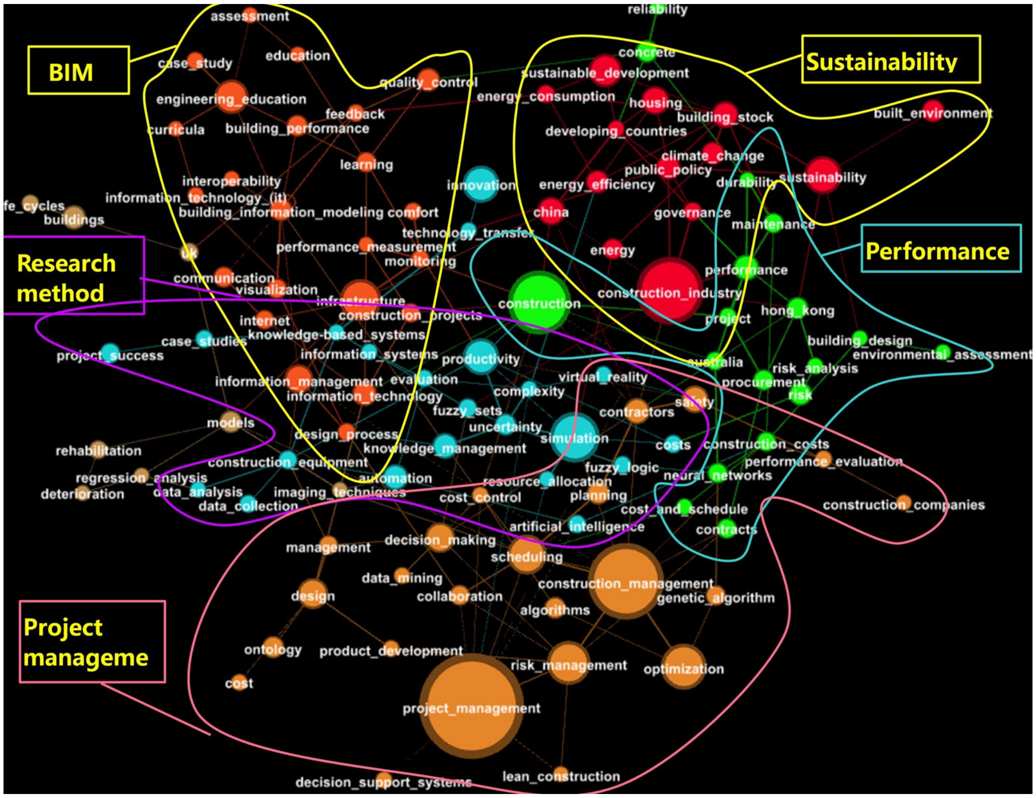

Figure 1 The Topic Network and Subgroups of CEM

The findings indicate that CEM research mainly contain five components of topics, including *project management, performance, research methods, BIM* and *sustainability*.

(1) Project management

The orange nodes in Figure 1 denote the topics of project management (Table 1), including certain topics in the management processes in construction projects , such as performance related topics incorporating *cost control, performance evaluation, planning and scheduling*; and some other topics such as *decision making, collaboration, design, risk management, lean construction, optimization, contractor* and *safety*.

Table 1 Topics of Project management

Topic	Fre	Year	Topic	Fre	Year
project management	463	1991	ontology	58	2004
construction management	321	1991	collaboration	51	2000
optimization	174	1994	management	45	1994
risk management	163	1991	performance evaluation	38	1991
scheduling	144	1994	decision support systems	35	1996
genetic algorithm	108	1998	cost control	34	2002
design	99	1993	product development	34	1991
decision making	97	1991	cost	34	2000
contractors	94	1995	data mining	33	1999
safety	84	1994	construction companies	29	2002
algorithms	64	2002	lean construction	29	2003

planning	63	1997	

Note: Year means the year that the corresponding topic firstly appear.

(2) Sustainability

The red nodes in Figure 1 denote the topics of sustainability (Table 2). This group incorporates environment issues, such as *sustainability, sustainable development* and *climate change*; public issues, such as China, *public policy,* and *governance*; and energy concerns, such as energy, energy efficiency and energy consumption.

Table 2 Topics of Sustainability

Topic	Fre	Year
Sustainability	135	1998
Sustainable development	113	1996
Housing	89	1994
China	89	1994
Building stock	89	1999
Public policy	72	1997
Energy efficiency	65	1996
Climate change	64	2001
Built environment	62	1996
Energy	54	1997
Governance	43	2000
Energy consumption	41	1994
Developing countries	34	1994

(3) BIM（building information modeling）

The pink nodes on the right of sustainability group, denote the topics of BIM (Table 3). The topics in this group could be further divided into three categories. First, there are some technical components related to BIM, including *information technology, visualization, monitoring* and *internet*. Second category focuses on BIM adoption and its benefits, such as *information management, building performance, quality control, assessment, interoperability* and *performance measurement*, *engineering education, communication, learning,* and *design process*.

Table 3 Topics of BIM

Topic	Count	Year	Topic	Count	Year
Building information modeling	206	2004	Construction projects	46	1996
Infrastructure	155	1996	Monitoring	42	1994
Engineering education	117	1991	Design process	40	1999
Information technology	106	1991	Case study	39	1991
Information management	92	1999	Feedback	38	1996
Quality control	66	1994	Comfort	38	1994
Building performance	57	1997	Assessment	31	1996
Communication	54	2001	Interoperability	30	2004
Learning	52	2000	Curricula	29	2002
Visualization	50	1997	Education	29	1999
Internet	46	2001	Performance measurement	29	2000

(4) Research Methods

The blue nodes, in the central part of Figure 1, represent the topics related to research methods (Table 4). Most of the nodes have intensive links with other groups, indicating that the topics in this group have a relatively widen applications. The topics include *simulation, automation, knowledge management, project success, artificial intelligence, fuzzy logic, data collection, case studies, virtual reality, complexity, information systems* and *data analysis.*

Table 4 Topics of Research Methods

Topic	Count	Year	Topic	Count	Year
Simulation	195	1995	Case studies	40	1995
Innovation	131	1995	Evaluation	37	2001
Productivity	128	1991	Virtual reality	35	2000
Automation	76	1999	Complexity	33	1998
Knowledge management	75	1997	Information systems	33	1994
Project success	51	1997	Data analysis	32	1991
Construction equipment	44	1995	Resource allocation	31	1996
Artificial intelligence	42	2000	Knowledge-based systems	31	1995
Costs	41	1994	Technology transfer	30	1995
Fuzzy logic	40	2000	Uncertainty	59	1996
Data collection	40	2003	Fuzzy sets	49	2001

(5) Performance

This group incorporates topics of performance in construction projects (Table 5), such as *construction, performance, risk, concrete, procurement, maintenance, contracts, construction costs, cost and schedule, risk analysis* and *reliability.*

Table 5 Topics of Performance

Topic	Count	Year	Topic	Count	Year
Construction	261	1994	Construction costs	47	2002
Performance	77	1994	Australia	44	1995
Risk	75	1999	Project	37	1998
Concrete	60	1996	Cost and schedule	33	2008
Hong kong	57	1995	Risk analysis	33	1991
Procurement	53	1998	Reliability	30	1999
Maintenance	50	1994	Building design	30	1996
Neural networks	50	1995	Environmental assessment	30	1999
Contracts	49	1995	Durability	29	1994

DISCUSSIONS AND CONCLUSION

Manual reviewing the literature is a complex process, and it is difficult to indicate the hidden connections between topics underlying the texts. Such a manual approach wastes time and energy, and more important, would lead to inevitable personal biases. Unlike the traditional approach in review works, this study explores the classic topics of CEM and maps their structure by using quantitative methods and SNA techniques from a large amount of data, providing relatively unbiased results.

The findings illustrate that *project management, sustainability, BIM* and *performance* are the main groups of CEM research, with distinguished focuses within groups and less connections between groups. Moreover, CEM research has applied a wide range of methods and models, with intensive connections with the main groups.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation of China (NSFC) (No. 71390522, NO. 71671053, NO. 71271065, NO. 71771067). The work described in this paper was also funded by the National Key R&D Program of China (No. 2016YFC0701800 and No. 2016YFC0701808).

REFERENCES

Chan, A.P.C., Chan, D.W.M., Chiang, Y.H., Tang, B.S., Chan, E.H.W., Ho, K.S.K., 2004. Exploring critical success factors for partnering in construction projects. Journal of Construction Engineering and Management-ASCE, 130(2), 188-198.

Floricel, S., Michela, J.L., Piperca, S., 2016. Complexity, uncertainty-reduction strategies, and project performance. International Journal of Project Management, 34(7), 1360-1383.

Flyvbjerg, B., 2014. What You Should Know About Megaprojects and Why: An Overview. Project Management Journal, 45(2), 6-19.

Foray, D., Lundvall, B., 1998. The knowledge-based economy: from the economics of knowledge to the learning economy. The economic impact of knowledge, 115-121.

Laurance, W.F., Clements, G.R., Sloan, S., O/'Connell, C.S., Mueller, N.D., Goosem, M., Venter, O., Edwards, D.P., Phalan, B., Balmford, A., Van Der Ree, R., Arrea, I.B., 2014. A global strategy for road building. Nature, 513(7517), 229-232.

Lin, G.B., Shen, Q.P., 2007. Measuring the performance of value management studies in construction: Critical review. Journal of Management in Engineering, 23(1), 2-9.

Meho, L.I., Yang, K., 2007. Impact of data sources on citation Topics and rankings of LIS faculty: Web of science versus scopus and google scholar. Journal of the American Society for Information Science and Technology, 58(13), 2105-2125.

Newman, M.E.J., 2004. Analysis of weighted networks. Physical Review E, 70(5).

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America, 103(23), 8577-8582.

Price, D.d.S., 1976. A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science and Technology, 27, 292-306.

Silva, F.N., Amancio, D.R., Bardosova, M., Costa, L.D., Oliveira, O.N., 2016. Using network science and text analytics to produce surveys in a scientific topic. Journal of Informetrics, 10(2), 487-502.

Small, H., 1973. Co-citation in the scientific literature: A new measure of the relationship between two documents. Journal of the American Society for

Information Science, 24(4), 265-269.
Wing, C.K., 1997. The ranking of construction management journals. Construction Management and Economics, 15(4), 387-398.
Zhu, L., Liu, X.T., He, S., Shi, J., Pang, M., 2015. Keywords co-occurrence mapping knowledge domain research base on the theory of Big Data in oil and gas industry. Scientometrics, 105(1), 249-260.