Advanced Engineering Informatics 34 (2017) 17–35

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

Long-term knowledge evolution modeling for empirical engineering knowledge

Xinyu Li [a], Zuhua Jiang [a,*], Bo Song [a,b], Lijun Liu [a]

[a] Department of Industrial Engineering and Management, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, PR China
[b] China Institute of FTZ Supply Chain, Shanghai Maritime University, 1550 Haigang Ave., Shanghai 201306, PR China

ARTICLE INFO

Article history:
Received 20 March 2016
Received in revised form 28 June 2017
Accepted 4 August 2017
Available online 17 August 2017

Keywords:
Knowledge evolution
Empirical engineering knowledge (EEK)
Evolution model
Knowledge representation
Data visualization

ABSTRACT

In this era of knowledge economy, appropriate management of the rapidly evolving knowledge is a real and urgent issue for factories and enterprises, in order to maintain the competitive edges. However, facing the onerous analysis required for understanding the long-term knowledge evolution, especially the evolving of empirical knowledge in the engineering field, effective and comprehensive modeling methods for knowledge evolution are absent. In this paper, a novel knowledge evolution modeling method is proposed for portraying the long-term evolution of empirical engineering knowledge (EEK) and assisting engineers in comprehending the evolving history. Three phases, EEK elicitation and formalization, EEK networks foundation, and family-tree evolution model construction, are included in the modeling method. This method is developed using natural language processing, semantic similarity calculation, fuzzy neural network prediction, clustering algorithm, and latent topic extraction techniques. To evaluate the performance of the proposed modeling method, an evolution model of empirical knowledge in computer-aided design (CAD) is constructed and then verified. Experimental results show that the proposed method outperforms the former approaches in feasibility and effectiveness, and hence opens up a better way of further understanding the long-term evolution course of EEK.

1. Introduction

In this era of knowledge economy, knowledge was fast maturing and mutating, driven by plenty of fresh concepts, techniques, methodologies, experiences and activities. The Internet and network technology also endowed knowledge with the characteristics of quick update, wide transmission and high interdisciplinary, which further promote the renewal of knowledge [1]. Under this situation, appropriate management of this rapid evolving knowledge is a significant task for the factories and enterprises, since it is the core to maintain their competitive edges in creativity and adaptability [2]. It also aroused an inevitable research question in the field of knowledge management: how the knowledge evolves over a long period of time? Since 1950s, researchers have realized that the networks constructed with the relations among academic research outputs could distinctly reveal the growth and inheritance of knowledge in some disciplines, hence solving the above question with some preliminary findings. Citation network for academic papers was firstly proposed for discovering the course of development [3]. It then expanded and enriched to co-word networks in the following researches [4]. To establish and assess the networks, metrology method is widely adopted for measuring the research development in different aspects: authors, research groups, countries, keywords, journals, etc. With the networks and metrology methods, the evolution of the knowledge presented and reviewed in academic literatures are objectively and quantitatively modeled [5–13].

However, in practice, modeling the evolution of academic knowledge is far from enough. Both in factory and enterprise, it's urgent to model the evolution of engineering knowledge, especially the evolution of empirical engineering knowledge (EEK). EEK is deduced and concluded from technical designs, decision-making or other engineering activities, which provides valuable instructions for the success and lessons from the failures. Having concluded and integrated in numerous relevant literatures [14,15,26–30], EEK can be defined as *a specific technical know-how about solving an engineering problem, which is a consequence of probable association and extension of engineering concepts under specific constraints of engineering scenarios, deduced and concluded from repeated observations, practices and communications of engineering technicians in long-term engineering activities.*

* Corresponding author.
E-mail address: Zhjiang@sjtu.edu.cn (Z. Jiang).

http://dx.doi.org/10.1016/j.aei.2017.08.001
1474-0346/© 2017 Published by Elsevier Ltd.

Unfortunately, two aspects impede the direct application of co-word network and metrology method in modeling EEK evolution: the "randomness" [14–15] and "tacitness" [16] of EEK. Except for this traditional methodology, several innovative modeling methods [17–24] have been designed in recent researches. They used the theories and tools in complex network analysis (CNA), or put forward some specific evolution modes. However, most of them either fail to consider the innate semantic and logical characteristics of EEK comprehensively, or lack the capability of portraying engineering domain or subdomain-level evolution courses. Such defects become more critical, as they analyze a proliferating engineering field with a huge amount of fast updating information and maturing knowledge in a long-term (such as one decade or more). Therefore, a customized and long-term-capable method for filtering colossal information and mining evolution courses of EEK shall be our top priority.

To achieve the object, this paper proposes a novel long-term knowledge evolution modeling method to reveal the courses of EEK evolution over several overlapping time periods. Based on the elicitation and formalization of EEK, the proposed method adopts semantic similarity calculation and fuzzy neural network prediction for constructing the EEK networks—a kind of network enlightened by the co-word network. Also, clustering algorithm and latent topic allocation are used to quantitatively discover the inheritance relationship in engineering fields. Presenting how EEK clusters evolve in a long time with a family-tree model, it will be easy to identify and visualize the courses of knowledge evolution and discover some interesting trends and patterns, thereby obtaining a better understanding of EEK evolution in a long-term.

The remainder of this paper is structured as follows. Section 2 uses some related works to introduce the empirical engineering knowledge (EEK) and its evolution. Additionally, some recent knowledge evolution modeling approaches are briefly analyzed in this section. Section 3 designs the general framework of the proposed evolution modeling method. The elicitation and formalization of EEK are illustrated in Section 4. Sections 5 and 6 detail the foundation of EEK networks and the construction of knowledge evolution model. The example of using the proposed method to model the evolution of EEKs originated from computer-aided design (CAD) missions, including the comparison and discussion with former works, is presented in Section 7. The last section concludes the paper with some possible improvements.

2. Related works

2.1. Representation of empirical engineering knowledge (EEK)

In factories and enterprises, EEK is of significant value for innovative design and decision-making process, and a good management of EEK will greatly promote the processes. As the first and foremost step in EEK management, a proper representation of EEK can largely facilitate the subsequent EEK acquisition, accumulation and reuse. Chan [25] firstly proposed that empirical engineering knowledge could be characterized by some discrete and standard attributes. Based on this idea, several scholars concentrated on EEK representation [14,15,26–30], and a concise summary of EEK representation in their works is listed in Table 1.

As summarized in Table 1, all researchers above regarded the core contents of EEK are *Problem*, *Context* and *Solution*. To comprehensively represent the capacity of EEK and relevance among EEKs, they also considered some more detailed attributes, such as *Effect*, *Contributors* and *Relevance*. Unfortunately, none of above researchers considers the timeliness of EEK in representation. In some engineering fields, such as electronic device design, advanced material manufacturing or software engineering, knowledge is updating frequently. The replacement of some obsolete empirical knowledge by some experience about new methods and equipment occurs rapidly, widely and deeply. Under this condition, it is hard for the existing representations to reflect the dynamics of EEK and portray its life cycle, since they are only designed for representing just one steady state of EEK. Therefore, *Time* should be another important attribute in EEK representation, serving as a hint for demonstrating the evolution of EEK.

2.2. Mechanisms of EEK evolution

The theory of evolution was initially designed for understanding and explaining the development of complex biological systems by Charles Darwin in 1842. Despite controversies regarding to this theory, it is widely adopted in numerous studies. As an analogy, in knowledge evolution, the mutation and selection of concepts, terminologies and approaches, are the fundamental causes of the knowledge evolution course that directs from specialty to generality, from vagueness to clarity, and from abstract to concreteness [31]. Basically, there are two knowledge evolution mechanisms accepted commonly by researchers: Darwinism knowledge evolution and Lamarckism knowledge evolution [22,24,32–36]. Having those relevant literatures reviewed, two evolutional mechanisms are compared in Table 2.

For the evolution of EEK, both mechanisms work simultaneously. On one hand, since the EEK is oriented from the solving of engineering problems, the necessary conditions of Darwinism knowledge evolution are nurtured. Along with the repeated cycles of utilization and update in various engineering situations, EEK is gradually developing. On the other hand, triggered by disruptive technological innovation, fundamental theories, tools or approaches varied, namely, paradigm shifts happened [36,37]. During the process of paradigm shift, the core contents of EEKs are partly or totally changed, and Lamarckism knowledge evolution generated. Therefore, in the research of long-term EEK evolution, characteristics of both two evolutional mechanisms should be considered. It's essential to perceive the coexistence of gradual and radical change in the evolution courses of EEK. In the proposed modeling method, they will be considered by the well-designed time windows.

Table 1
Considered attributes in EEK representation.

Representative Works	Problem/ Goal	Requirement/Context/ Cause	Conclusion/Solution/ Response	Credibility/ Effect	Contributors/Participant/ Individual	Relevance/ Reference
Chen [15]	√	√	√			
D'Eredita [26]	√	√	√			
Argote [27]	√	√	√		√	
Foguem [28–29]	√	√	√	√		
Zhang [30]	√	√	√	√		√
Liu [14]	√	√	√	√	√	√