Safety Science 109 (2018) 46–56

Contents lists available at ScienceDirect

Safety Science

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

Big-data-driven safety decision-making: A conceptual framework and its influencing factors

Lang Huang*, Chao Wu, Bing Wang, Qiumei Ouyang

School of Resources and Safety Engineering, Central South University, Changsha, Hunan 410083, P R China
Safety & Security Theory Innovation and Promotion Center, Central South University, Changsha, Hunan 410083, PR China

ARTICLE INFO

Keywords:
Big data (BD)
Safety big data (SBD)
Safety decision-making (SDM)
Safety insight (SI)
Data-driven

ABSTRACT

Safety data and information are the most valuable assets for organizations' safety decision-making (SDM), especially in the era of big data (BD). In this study, a conceptual framework for SDM based on BD, known as BD-driven SDM, was developed and its detailed structure and elements as well as strategies were presented. Other theoretical and practical contributions include: (a) the description of the meta-process and interdisciplinary research area of BD-driven SDM, (b) the design of six types of general analytics and five types of special analytics for SBD mining according to different requirements of safety management applications, (c) the analysis of influencing factors of BD-driven SDM, and (d) the discussion of advantages and limitations in this research as well as suggestions for future research. The results obtained from this study are of important implications for research and practice on BD-driven SDM.

1. Introduction

Safety data and information are the most valuable assets for organizations' safety decision-making (SDM), especially in the era of big data (BD) (Wang and Wu, 2017). Access to large-scale, fast-moving, complex streams of safety data-sets has the potential to fundamentally transform the way organizations make their safety decisions. Accordingly, the BD support for SDM at all levels of organizations, along with the way them are organized, is becoming increasingly critical.

1.1. Problems of traditional SDM methods

In the safety management domain, SDM is the basic activity of safety managers, and has been an important and thoroughly covered research topic throughout the years. According to research literature (Hayes, 2012; Cha and Ellingwood, 2013; Yu et al., 2013; Talarico and Reniers, 2016; Ma et al., 2016; Shafiee and Animah, 2017; Simanaviciene et al., 2014; Insua et al., 2016; and Hersing, 2017), there are four types of traditional SDM methods, namely, manual inspection, statistical tools, expert systems and mathematical modeling methods.

(a) The manual inspection methods mainly rely on the safety decision-makers' senses and experiences, and the gut-feeling and intuition-based heuristics are also involved. Thus, this type of method is faced with uncertainties caused by subjective factors. Moreover, it is time-consuming and labor-consuming.
(b) The expert systems methods tend to have a relatively lower level of reliability and correctness of SDM, because the safety decisions are usually generated using relevant expertise that has been accumulated for a long time, and only a limited amount of safety information is used. Additionally, it is difficult to unify the various safety decisions acquired by different experts, thus affecting the efficiency of decision-making and actions.
(c) Performance of the statistical methods and tools in SDM is weak in terms of applicability and reliability due to the following problems: the small size of collected safety data that makes it difficult to extract valid safety information, the scattered distribution and invalid integration of safety data sets, lack of efficient safety data analysis tools and lack of timely collection and updating of important safety data.
(d) The mathematical modeling methods have several disadvantages. Firstly, it is difficult to establish accurate mathematical models due to the increasingly complexity of social-technical system safety problems. Secondly, the modeling process often involves human assumptions, thus reducing the actual complexity of the safety theoretical models. Thirdly, it requires a lot of computing time and execution time, which makes it difficult for these methods to meet the real-time and dynamic demands of actual SDM.

Based on the analysis of the traditional SDM methods, there is an

* Corresponding author at: School of Resources and Safety Engineering, Central South University, Changsha, Hunan 410083, PR China.
E-mail addresses: 18971185983@189.cn (L. Huang), wuchao@csu.edu.cn (C. Wu), wangbing187717@163.com (B. Wang), ouyangqiumei0893@csu.edu.cn (Q. Ouyang).

https://doi.org/10.1016/j.ssci.2018.05.012
Received 17 January 2018; Received in revised form 22 April 2018; Accepted 16 May 2018

urgent need to develop new methods to address the above problems in the area of BD.

1.2. The propose of BD-driven SDM

BD refers to extremely large amounts of structured, semi-structured or unstructured data continuously generated from diversified sources, which has an impact on decision-making through mining insightful information from rambling data (Gandomi and Haider, 2015; Sheng et al., 2017; Cheng et al., 2018). In the safety management domain, with the rapid development of emerging information technologies, a large amount of multi-source and heterogeneous safety-related data – known as SBD – has been accumulated in almost every aspect of people's lives, where intrinsic details and patterns of hidden knowledge should be extracted and utilized. From the perspective of safety decision makers, the significance of SBD lies in its ability to provide safety information and safety knowledge of value, upon which to base safety decisions. Therefore, to facilitate SDM, organizations need efficient methods to process large volumes of assorted SBD into meaningful SI. However, the "5V" features of SBD, namely, huge Volume, high Velocity, high Variety, low Veracity, and high Value (Sheng et al., 2017; Emani et al., 2015; Bello-Orgaz et al., 2016), make it difficult for traditional computing methods to effectively support the processing, analysis and computation of SBD. Accordingly, a challenge posed to today's organizations is how to identify methods to align SBD that could help them make better-informed safety decisions.

In recent years, more advanced analytics and platforms have been invented and employed to handle BD, which can outperform traditional techniques in every aspect of data management and analytics (Kambatla et al., 2014; Tiwari et al., 2018; Gupta and George, 2016). BD analytics and technologies enable users to capture, store, and analyze huge amounts of data of an unprecedented breadth, depth and scale, obtaining meaningful insights from multiple sources both inside and outside the organization (Miah et al., 2016; Purarjomandlangrudi et al., 2014; Pramanik et al., 2017; Wang et al., 2016b). Therefore, BD analytics have been playing a central role in many decision-making domains, such as supply chain management, smart health, fault diagnosis, business analysis, product development, smart grids and smart energy management. Extant studies have demonstrated that substantial value and significance in finding methods to promote emerging interdisciplinary research, to better perceive the present, to better predict the future and to lower management risks and improve operation efficiency, can be attained by organizations through taking effective decisions based on BD analytics (Frederiksen, 2009; Jin et al., 2015; Zhong et al., 2016; Gunasekaran et al., 2016; Addo-Tenkorang and Helo, 2016; Munshi and Mohamed, 2017; Manco et al., 2017). The above research has proved that effective and successful organizational management lies in the ability to apply data analytics to extracting valuable insights from BD to support decision-making.

In light of the great significance of BD-driven decision-making, BD-based methods may address the challenges faced with traditional SDM, for BD analytics and data mining technologies can be used to conduct an in-depth analysis into SBD to discover valuable SI, and, thus, to make more felicitous safety decisions. Actually, researchers and practitioners, such as Ouyang et al. (2017), Huang et al. (2018), Shi and Abdel-Aty (2015), Guo et al. (2016) and Newell and Marabelli (2015), have adopted BD thinking, applied it to the safety domain and proved its theoretical and practical benefits in safety research. However, there are still many challenges to the implementation of SBD, of which the primary and urgent ones include: (a) how to effectively and completely detect, collect, store and manage SBD; (b) how to efficiently analyze and mine multi-source heterogeneous SBD; (c) how to turn huge volumes of SBD into valuable and workable safety information; (d) how to use SBD to support more effective and efficient SDM; and (e) how to identify the factors influencing SDM quality from the BD perspective.

1.3. The aim of this study

In light of the challenges above, the priority for organizations is to develop an effective and efficient data-mining process to extract useful but hidden SI from SBD. This paper aims to research how BD analytics can be integrated into the SDM process to enhance decision-making and provide valuable SI for safety decision makers. To this end, a conceptual framework for SDM based on BD, known as BD-driven SDM, was developed, and its influencing factors were discussed. By adopting this framework, safety decision makers could be able to enhance the quality of the decision-making process, and thus the quality of the safety decisions. Notably, this study mainly illustrates the ideas and methods from a theoretical perspective.

The remainder of this paper is structured as follows. Section 2 outlines the theoretical bases for the proposed framework. Section 3 describes the developed BD-driven SDM framework. Section 4 investigates the factors influencing safety decision-making. Section 5 discusses the advantages and limitations, and provides suggestions. Finally, Section 6 summarizes the conclusions.

2. Theoretical foundations

Interdisciplinary research offers diverse motivations for new research questions and theoretical perspectives relevant to all the phases of the research process – from problem and theory formulation to research model and method selection, all the way to data analysis and interpretation of the results (Chang et al., 2014). With the rapid development of scientific research, previously distinct disciplines have converged to integrate a variety of functions on demand, including interactions among the disciplines of marketing, economics and computer science, where fresh insights on consumer behavior and product design provide a new basis for competitive advantage, as well as interdisciplinary research areas, such as energy, social and information science, where a deep understanding of household energy consumption behavior can provide an effective way to improve energy efficiency and promote energy conservation (Zhou and Yang, 2016).

In the SDM domain, the assorted SBD describing the complex organizational safety mechanism that researchers would like to investigate cannot be adequately understood from a single disciplinary perspective. For example, a huge amount of data is available for the study of human unsafe behaviors in safety science, but it cannot be analyzed without data-mining and machine-learning methods of data science. Accordingly, an interdisciplinary approach is needed to deal with the emergence of SBD.

2.1. The meta-process of BD-driven SDM

SBD is first collected from different data sources with various qualities and then processed by various organizational entities, resulting in a flow or chain of activities which can be labeled as the "BD chain". In other words, the process of SDM consists of various safety data sources, several data-processing activities and various organizations. Accordingly, a deep understanding of the BD chain is necessary to developing the BD-driven SDM framework. For research clarity, a brief review of the steps of the BD chain can be seen in Table 1.

According to the fundamental and widely recognized information-processing model (Rowley, 2007), there is a data – information – knowledge – wisdom (DIKW) hierarchy. When it comes to BD-driven SDM, the DIKW hierarchy can be extended into a new one (Huang et al., 2017)), that is, the safety data – safety information – safety knowledge – safety science theory. Accordingly, in the present research, the process of BD-driven SDM can be broken down into four stages: data sensing, capturing and storing; data cleansing, integration, transformation and reduction; data mining and SI extracting; and SI interpretation and application. They are described as follows: