doi: 10.12011/1000-6788(2018)07-1877-11 中图法分类号: X91 文献标志码: A

大数据视阈下的系统安全理论建模范式变革

黄 浪, 吴 超, 王 秉

(1. 中南大学 资源与安全工程学院, 长沙 410083; 2. 中南大学 安全理论创新与促进研究中心, 长沙 410083)

摘 要 为明晰大数据视阈下的系统安全理论建模范式变革, 首先, 分析数据和信息在系统安全研究中的双重性质演变. 其次, 构建包括数据维、系统维和安全维的系统安全理论建模范式转变框架, 采用理论思辨法, 分析大数据视阈下现有系统安全理论建模面临的挑战与机遇. 最后, 在分析建模技术路径、建模逻辑主线和建模原理的基础上, 构建基于大数据的系统安全理论建模新范式, 并进行实例分析. 结果表明: 数据和信息在系统安全中具有主体和客体双重性质; 系统安全理论建模范式转变可分为 3 阶段: "经验 - 微系统 - 古典安全"、"小数据 - 中系统 - 近代安全" 和 "大数据 - 宏系统 - 大安全" 范式; 现有的 "小数据型"、"解释型" 等 9 类安全理论模型不能适应大数据时代的系统安全需求, 大数据给系统安全理论建模带来的机遇包括面向全体数据、突出预测性等 6 方面; 实例分析证明基于大数据的系统安全理论建模新范式具有可行性和科学性.

关键词 大数据; 安全理论模型; 系统安全; 安全数据; 安全信息

Perspectives on paradigm shift of system safety theoretical modeling in the era of big data

HUANG Lang, WU Chao, WANG Bing

(1. School of Resources and Safety Engineering, Central South University, Changsha 410083, China; 2. Safety & Security Theory Innovation and Promotion Center (STIPC) of CSU, Changsha 410083, China)

Abstract In order to understand the paradigm shift of system safety theoretical modeling based on big data, firstly, the double characteristics of data and information in system safety research was analyzed. Secondly, a framework for paradigm shift of system safety theoretical modeling, which consist of data dimension, system dimension, and safety dimension, was constructed. Then, using the theory of speculation, the challenges and opportunities of system safety theoretical modeling based on big data were analyzed. Lastly, based on the analysis of modeling technology path, modeling logic line and modeling principle, a new paradigm system safety theoretical modeling based on big data was constructed, and case analysis was carried out. Research results show that: the paradigm shift of system safety theoretical modeling can be classified into three stages, which are "experience-micro-classical safety", "small data-meso-modern safety", and "big data-macro-big safety". The existing nine categories of safety theoretical model cannot meet the safety and security requirements in the era of big data, the opportunities brought by big data can be analyzed from six sides. Case analysis proves that the new paradigm of system safety theoretical modeling based on big data is feasible and scientific.

Keywords big data; safety theoretical modeling; system safety; safety data; safety information

收稿日期: 2017-08-03

作者简介: 黄浪 (1990–), 男, 四川泸州人, 博士研究生, 研究方向: 安全科学基础理论、系统安全理论建模原理研究, E-mail: 18971185983@189.cn; 吴超 (1957–), 男, 广东揭阳人, 工学博士, 中南大学资源与安全工程学院教授、博士生导师, 兼任教育部高等学校安全科学与工程学科教学指导委员会委员、全国安全工程领域工程硕士培养协助组副组长、国家安全生产专家、公共安全科学技术学会理事等, 国务院政府津贴获得者, 研究方向: 安全科学与工程基础理论, E-mail: wuchao@mail.csu.edu.cn.

基金项目: 国家自然科学基金重点项目 (51534008)
Foundation item: Key Project of National Natural Science Foundation of China (51534008)
中文引用格式: 黄浪, 吴超, 王秉. 大数据视阈下的系统安全理论建模范式变革 [J]. 系统工程理论与实践, 2018, 38(7): 1877–1887.
英文引用格式: Huang L, Wu C, Wang B. Perspectives on paradigm shift of system safety theoretical modeling in the era of big data[J]. Systems Engineering — Theory & Practice, 2018, 38(7): 1877–1887.

1 引言

对于任何一个系统, 其安全科学核心问题是安全理论模型构建. 安全理论模型是系统安全各要素间逻辑关系和机制的抽象表述 [1], 是安全科学知识体系的基石, 是事故预防与控制的钥匙, 是构筑系统安全的指南. 构建安全理论模型, 把安全理论建模作为系统安全研究的一种手段与方法, 是人类在认识系统安全和塑造安全系统过程中的一大创造 [2]. 随着社会技术系统复杂性的提高, 尤其是进入大数据时代、工业 4.0 时代、人工智能时代以后, 系统复杂性与耦合性、数字化与智能化的快速提高使传统的安全理论模型不能满足复杂系统安全研究与实践需求.

数据是科学研究的重要基础, 大数据已经成为各行各业的研究热点, 《Nature》[3] 和《Science》[4] 分别于 2008 年、2011 年推出有关大数据的专刊. 尽管现在还存在大数据的出现是否推动了科学研究的第四范式 (即数据密集型科学研究) 的产生的质疑 [5], 但不可否认的是, 大数据作为人类认识世界的一种新方法与新工具, 在改变我们的生活、工作和思维方式的同时, 也对科研思维和科研方法产生深远影响, 产生了数据密集型和驱动型科研方法, 并在诸多领域得到广泛应用, 如智能农业 [6]、战略运输 [7]、供应链管理 [8]、智能决策 [9]、计算型社会科学 [10]、地球科学 [11]、疾病探测 [12] 等. 在安全科学领域, Huang[13]、欧阳秋梅 [14,15]、王秉 [16] 等对大数据应用于安全科学领域的基础原理进行了探讨, Shu[17]、Guo[18]、Shi[19] 等对大数据应用于过程安全分析、行为安全管理、交通安全等领域进行了实践研究. 上述无论是大数据应用于其他领域还是安全科学领域, 其基础性问题都是理论模型构建, Gil 等 [20] 提出大数据是一种新的建模和管理方法, 但还缺乏方法论和范式相关研究. 所谓范式是指一段时间内为科学共同体共同遵从的、用于指导科学研究的一组理论、准则和方法的总和.

因此, 本文将在大数据时代数据驱动的 “科研信息化” (e-science) 背景下, 探讨传统系统安全理论建模所遇到的挑战与机遇, 并分析基于大数据的系统安全理论建模范式变革, 以期为后续大数据应用安全科学领域或其他学科领域提供理论指导.

2 数据与信息在系统安全中的双重性质演变

安全理论建模主客体关系的本质是安全生产活动的主客体关系. 随着大数据、人工智能等研究的深入, 信息社会不断推进, 人类生存、生产与生活活动关系发生了重大改变. 宋守信等 [21] 以甬温线动车事故为例, 分析了信息社会背景下安全生产活动主客体关系的变化, 而大数据技术正是信息社会从 “量变” 到 “质变” 的关键推力, 信息化社会在大数据时代才算真正到来. 传统生产活动的主客体关系 (主体是人, 客体是物质实体) 决定了传统安全理论建模的主客体关系 (主体是人, 客体是物或机), 如经典的轨迹交叉模型认为人 (主体) 的不安全行为和物 (客体) 的不安全状态相交叉 (同一时间、同一空间发生) 必然导致事故发生. 进入大数据时代, 大数据、人工智能等技术的应用从根本上改变了复杂社会技术系统的人、机、环安全生产活动关系, 这必然对安全理论研究与实践产生变革性影响. 因此, 探讨大数据背景下安全生产活动主客体关系的演变, 对探析系统安全理论建模在大数据时代的机遇与挑战至关重要.

复杂社会技术系统客体组成元素的构成和演变远远超出了人的反应速度, 以数据和信息为驱动的人工智能扩展增强了人的认识、分析和决策能力, 在一定情境下同时也担当着安全生产活动主体的角色, 在某些领域开始出现 “无人化”, 如: 自动驾驶系统、自动飞行系统、高速列车控制系统等, 实现对信息的自动采集、处理、决策和执行等功能于一体. 以 “人 - 车” 系统为例: 在蒸汽机时代, 机车运行速度慢, 依靠机车司机和调度人员的 “目视判断”、“手动操作” 就能实现机车系统安全运转. 进入电力时代, 机车大幅度提速, 仅靠机车司机和调度人员的感知觉系统和较低层次的自动化水平已经不能满足机车系统安全运转要求, 进而出现了机车自身的软件控制系统和机车运转的软件调度系统. 在如今的高速铁路时代, 必须借助庞大的数据系统、信息系统与软件系统的智能控制, 才能实现机车的安全运转. 虽然人依然是数据、信息和软件的主要控制者, 但是人们安全信息的获取、安全信息的分析、安全预测与决策、安全信息利用等安全行为需要数据、信息和软件的智能支撑. 换言之, 以数据、信息为驱动的人工智能系统延伸了人类的智能, 在安全生产活动中即具有主体性质又具有客体性质.

由上分析可知, 传统安全理论建模将数据和信息系统当做人 - 机系统中的机来对待, 忽视了数据和信息系统所具有的类似人的能动性. 实际上, 进入大数据时代, 安全活动的主体已经超出了生物人的范畴, 出现了人脑与电脑组合的控制系统 [21]. 换言之, 安全生产活动的主体已经从单纯的人变成了由人和以数据信息为驱动的智能系统构成的整体, 数据和信息系统在安全生产活动中既是客体又是主体, 具有主客体双重性质 (见图 1). 在进行系统安全建模时, 必须考虑数据和信息在安全生产活动中的主体角色. 此外, 由于数据和信息技术的限制, 传统人 - 机系统分析模式下的人 - 机界面属于物理层面的人 - 机界面 (人和机直接接触), 重点关注的是处于人 - 机界面的单人、单机, 人 - 机交互受到时空限制. 但是在大数据时代, 随着数字和信息技术的快速发展, 已经由物理层面的人 - 机界面发展为数字化、信息化的人 - 机界面, 在该人 - 机交互模式下, 系统安全分析需要关注多人 (人群)、多机 (机群), 人 - 机交互不受时空限制, 数字和信息已经成为人 - 机交互的核心纽带, 数字技术和信息技术成为系统安全分析与控制的关键手段.

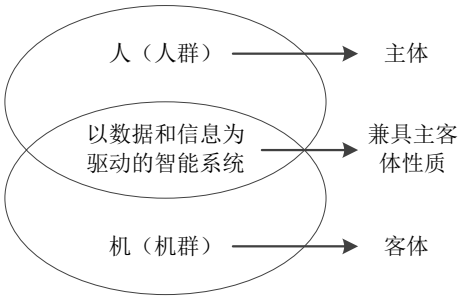

图 1　系统安全中数据和信息的主客体双重性质

3　大数据时代系统安全理论建模的挑战与机遇

3.1　数据驱动的系统安全理论建模范式转变

基于 “1 数据和信息在系统安全中的双重性质演变” 分析, 可构建系统安全理论建模范式转变 “数据 - 系统 - 安全” 3 维框架 (见图 2). 解析如下: ①数据维为, 根据安全数据技术自身的发展, 以及人们对数据在安全生产活动中的角色演变的认识, 可将数据维划分经验型认识阶段、小数据时代和大数据时代. ②系统维, 根据人们对安全的认识范围 (系统) 的扩大, 从系统粒度视角, 可将系统维划分为微系统、中系统和宏系统 3 阶段, 例如生产车间、工段属于微系统, 生产企业或组织属于中系统, 生产组织所属的经济社会系统属于宏系统. ③安全维, 根据刘潜 [22]、Stoop[23] 等对安全科学研究沿革的论述, 将系统安全研究划分为 3 个阶段, 分别是农业社会时期的古典安全范式、工业社会时期的近代安全范式和信息社会的大安全范式.

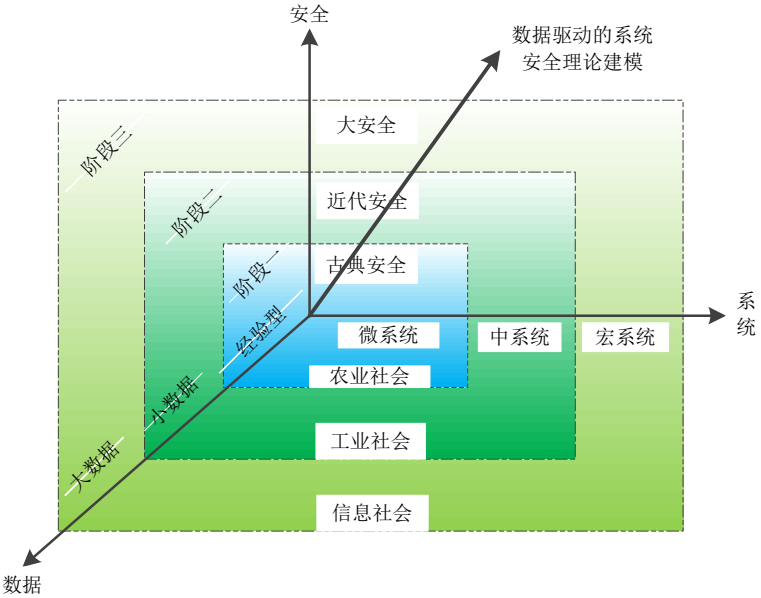

图 2　数据驱动下的系统安全理论建模范式转变 3 维模型

根据上述 3 个维度的划分, 可将数据驱动的系统安全研究分为 3 范式: ①范式一处于农业社会时期, 该时期人们只关注来自自然环境的危险, 主要靠日积月累的经验面对危险, 所关注的也只是个体所在的局部范围 (微系统), 但并没有系统安全研究, 可称之为 “经验型 - 微系统 - 古典安全” 范式. ②范式二处于工业社会时期, 人们面对的危险主要来自新技术与新工业, 系统安全研究开始萌芽与发展, 开始重视安全数据的重要性, 并通过数据统计得出系列安全法则指导安全生产 (如海因里希法则), 安全研究从只重视微系统层面的单人或单机, 到重视人 - 机交互, 再到采取系统思维考虑整个微系统安全, 随着系统安全研究的深入, 又由微系统安全研究转向了中系统安全研究, 该阶段的安全研究可称之为 “小数据 - 中系统 - 近代安全” 范式. ③范式三处于信息社会时期, 前已述及, 大数据技术使信息社会 “量变” 到 “质变”, 人们面对的安全现象涉及复杂的社会技术系统 (宏系统), 大安全观进入研究视野, 该阶段的安全研究可称之为 “大数据 - 宏系统 - 大安全” 范式, 或 “大数据 - 大系统 - 大安全” 范式. 需指出的是, 图 2 中 3 个阶段并没有严格的划分界限, 相互之间存在交叉.

在 “大数据 - 大系统 - 大安全” 范式下, 和诸多技术创新和思维革新一样, 大数据应用于系统安全理论建模的驱动力主要是 “挑战 - 机遇” 机制或 “拉 - 推” 机制 (pull-push mechanism). 拉力 (挑战): 为了实现既定目标需要新技术与新方法; 推力 (机遇): 因为新技术使人或组织能够实现更高更新的目标. 解释如下:

3.2 现有安全理论模型面临的挑战 (拉力)

1) 在现有的 “中小系统型” 安全理论模型方面. 从文献 [24] 的分析可以看出, 目前的安全理论模型主要集中于微观系统和中观系统, 但是随着社会技术系统 (sociotechnical system) 复杂性、耦合性和智能性的快速提高, “中小系统型” 安全理论模型将不能满足复杂系统安全分析需求. 而这些变化都是以数据和信息驱动为基础的, 复杂系统安全对数据和信息的依赖性更强; 同时复杂社会技术系统将产生大量安全数据, 由于传统数据统计和分析方法的限制, 传统的事故致因建模可能忽视或简化了一些致因因素, 将不能适应安全数据的指数型增长. 此外, 根据 Harvey 和 Stanton[25] 从 10 个方面论述的复杂系统安全 “10 大挑战”, 传统的系统安全建模思维和方法将不能满足复杂系统安全需求, 未来复杂系统安全研究需要以现代计算机技术和信息技术为基础的新技术与新思维.

2) 在现有的 “还原型” 安全理论模型方面. 安全理论建模方法论从宏观来说主要由整体论方法和还原论方法论构成. 基于还原论的建模方法预设系统安全问题的 “某一侧面”、针对某个问题来获取安全数据. 基于整体论的建模方法不分解系统, 把系统视作一个整体, 主要从系统的输入输出判断系统结构和功能. 传统安全理论建模主要采用还原论方法, 把复杂、多样、多变的安全现象首先通过还原论还原为某个逻辑基点, 即将系统不断地分解, 找出系统的构成组份及其内部机制 (可能遗漏一些关键要素), 针对各个逻辑基点和系统组份构建安全理论模型, 以解释系统的行为和功能. 从事故致因模型的发展与演变主线 (“点型事故模型”(以人或机为中心)→“线型事故模型”(链式事故模型)→“面型事故模型”(轨迹交叉类事故模型)→“体型事故模型”(系统事故模型)), 和从不同系统粒度视角对安全理论模型的划分 (微系统安全模型、中系统安全模型和宏系统安全模型) 可以看出, 随着安全科学的发展, 安全理论建模已经进入 “系统范式”, 需要系统论 (或整体论) 思想与方法, 从系统安全本源出发构建系统安全理论模型.

3) 在现有的 “小数据型” 安全理论模型方面. 在传统安全理论建模范式下, 由于数据采集、数据储存、数据分析等技术的限制, 可获得的数据比较少、冗余数据少、数据结构和类型单一, 可应用抽样统计方法, 通过分析少量安全数据得出系统安全特征, 用尽可能少的数据来发掘和表征尽可能重大的发现. 如海因里希冰山模型或海因里希法则 (通过对 55 万件机械事故的调查得出)、事故频发倾向论 (通过泊松分布、偏倚分布、非均等分布进行统计分析得出)、人因调查工具模型 (HFIT, 通过分析 18 个事故报告得出) 等等, 该类安全理论模型极大程度地推进了安全科学原理研究, 但只能反映特定统计样本的事故原理. 在传统安全统计建模范式下, “55 万” 样本统计结果已经令人信服, 但是在大数据视阈下, “55 万” 也属于小样本. 换言之, 基于大数据建模范式, 海因里希冰山模型等 “小数据型” 安全理论模型的准确性和适用性将受到冲击与挑战.

4) 在现有的 “静态型” 安全理论模型方面. 传统安全理论建模假设系统处于某个特定的时空位置, 系统的结构 (组成元素、元素之间的关联关系) 和功能都是确定的. 实际上, 无论是微系统、中系统还是宏系统, 其结构是随着时空变化而变化, 系统功能也随着时空变化, 即便是相同的系统结构在不同的时空位置系统功能

也可能不一样. 这也是导致现有安全理论模型滞后于科学技术发展 (或安全科学发展落后于科学技术发展) 的原因之一. 因此, 不同尺度层面的微观系统、中观系统和宏观系统结构和功能都处于不断变化之中, 传统的 “静态型” 安全理论建模方法不能适应复杂系统动态变化.

5) 在现有的 “解释型” 安全理论模型方面. 解释已经发生的安全现象和预测将来可能发生的安全现象是复杂系统安全理论模型的 2 个主要功能. 所谓解释就是对已经发生的现象找出因果或相关关系来说明现象之间的规律或关系. 所谓预测就是已知一些现象, 通过因果或相关关系来预见未来即将发生的现象. 以事故致因模型为例, 现有的事故模型大都属于事故发生以后, 通过总结和提炼事故致因的共性规律而构建, 并随着新事故致因的出现不断丰富和完善已有事故模型. “解释型” 安全理论模型能够分析与解释事故致因, 也能根据模型制定事故预防策略, 但只限于分析与预防已经发生过的类似事故, 不能实现对没发生过的、未知的事故进行分析与预测. 这也是一般事故和较大事故能够有效预防预控, 而重特大事故频发的重要原因, 因为重特大事故总是孕育着新的事故致因原理. 因此, 亟需借助新思维范式开展 “预测型” 安全理论模型研究.

6) 在现有的 “因果型” 安全理论模型方面. 因果性分析是各门学科的核心, 尝试从事物之间的因果关系来捕捉事物之间的基本规律. 在传统科学研究中, 基于简单的小数据系统, 比较容易做到因果分析. 以事故致因模型为例, 因果分析法是事故分析的基本准则, 也是事故致因建模最重要的方法 (原因 $\rightarrow$ 结果、结果 $\rightarrow$ 原因或原因 $\leftrightarrow$ 结果), 即通过分析事故的原因 (直接原因、间接原因、基本原因、根本原因、根源原因等), 并理清这些事故致因之间的层次与逻辑关系, 构建事故致因模型. 该类建模方法适合于简单线性系统, 可以详细地研究每个数据之间的关联, 并从中找出它们之间的因果关系和微观规律. 但是复杂系统安全问题实际上是非线性问题, 复杂系统安全建模面临的数据量大、冗余数据多、数据结构繁杂, 由于非线性、不确定性、复杂性导致因果关系不能确定 (或不能快速、经济、有效的确定), 一味的追求因果关系将不利于快速准确的安全预测与安全决策.

7) 在现有的 “抽象型” 安全理论模型方面. 在系统事故仿真建模方面, 现有安全理论模型能够分析事故致因因素和致因关系, 但没有考虑事故生命周期的过程仿真建模 (即事故的孕育、发生、发展、扩大与消亡过程的仿真再现), 随着计算机技术、虚拟仿真技术的发展, 所构建的安全理论模型需要考虑为系统事故 (或安全) 的仿真建模提供理论基础. 系统安全的仿真建模可以实现对系统 “事故 $\leftrightarrow$ 风险 $\leftrightarrow$ 安全” 演变的仿真实验, 这能够极大程度地弥补安全科学理论研究 “思维实验” 缺乏检验和修正的缺陷. 但目前的系统安全理论建模主要基于 “思维实验”(这是和其他学科的仿真实验、技术实验、物理实验和工程实验最大不同), 只能通过安全理论模型进行思维层面的推演, 还不能实现系统安全动态演变的仿真模拟.

8) 在现有的 “简化型” 安全理论模型方面. 由于传统安全分析方法、安全思维模式的限制, 在安全理论建模时, 首先是对目标系统进行简化, 找出基于当下安全理论水平的主要因素、主要属性、主要矛盾和主要关系 (该过程本身还具有主观性), 对目标系统所处的状态、环境和条件进行分析比较, 做出合理的简化与假设, 以便能够运用已有的科学知识和科学工具, 用低层次事物和比较简单的模型去解释复杂系统安全问题. 该类模型只是基于系统的某一侧面与某一假设, 没有把系统固有的安全属性、安全矛盾和安全关联关系 100% 的描述出来, 可能遗漏了一些关键因素, 这对于复杂系统安全建模将是致命性缺陷.

9) 在现有的 “小综合小交叉型” 安全理论模型方面. 传统的安全理论建模研究与实践中, 由于思维与方法的限制, 导致研究者只专注于各自领域的安全问题、安全理论与安全方法, 这具有以下缺陷: ①理论层面, 从安全科学 “上游中游下游” 的学科范式来看, 下游各个行业学科的安全数据、安全信息与安全知识之间处于独立状态, 不能实现相互关联、融合与渗透, 阻碍产生共性安全科学理论 (上游), 进而不利于安全科学理论的发展与完善. ②实践层面, 各领域研究者只能从各自的专业视角去分析复杂系统安全问题, 但是复杂系统安全问题具有非线性和不确定性, 需要跨学科、跨领域合作. ③经济适用性层面, 各个学科、各领域安全数据的封闭, 会造成科研资源的巨大浪费. 由此, 复杂系统安全理论建模亟需安全科学研究者或其他学科研究者跳出自己所在领域、行业和学科所固有的思维定式, 从一个更广阔的系统视角去理解和抓住复杂系统安全理论建模的多维属性 [26]. 此外, 从学科属性来看, 安全科学是一门典型大综合大交叉学科, 理应需要跨学科、跨领域、跨部门的建模范式, 构建 “大综合大交叉型” 安全理论模型.

3.3 基于大数据的系统安全理论建模机遇 (推力)

相比于传统的安全理论建模方法, 大数据技术可为系统安全理论建模范式带来如下变革:

1) 促使系统安全理论建模面向全体数据和动态数据: 在小数据时代 [27], 由于采集、记录、存储和分析具有复杂、多样、多变等特性的安全数据的技术和能力有限, 准确分析海量安全数据、多样性安全数据 (半结构化和非结构化数据) 是一种挑战. 进入大数据时代, 安全数据已成为安全理论建模研究与实践的核心, 大数据技术使安全理论建模面向全体安全数据并提供具体可行的技术途径, 面对复杂系统安全, 可以充分利用大数据对研究对象实现实时与全面描述, 并对海量安全数据进行分析和处理, 从而发现复杂系统安全规律或本质.

2) 推动安全理论建模实现还原论与整体论的融贯: 基于大数据技术, 不是局部收集随机样本, 而是全维、多角度分析系统安全数据. 由于处理了所涉问题的全部数据, 这就让整体论中所说的全面、完整的把握研究对象就有了科学的表述并落实到了具体的数据. 而这些全部数据是由一个个具体的数据构成的, 因此还原论中的要素、部分也得到了科学的表述. 大数据技术的出现为系统的整体性和动态性分析提供了条件, 放弃还原论的分解建模, 代之以 “整体数据” 的分析, 承认对复杂问题无法建模, 而是直接从 “现实” 去寻找答案, 这可能是新的建模思路. 此外, 大数据可以将分解出来的各种碎片又重新组成一个网络, 再次回到整体. 因此, 大数据技术给安全理论建模为实现复杂性科学的还原论与整体论的辩证统一提供了具体的技术实现路径.

3) 促使系统安全理论建模突出相关关系: 相关性是指两个或两个以上变量的取值之间存在某种规律性, 可通俗理解为一个变量的变化有可能引起另一个变量产生相应的变化, 因果关系属于相关关系. 传统安全理论建模面对的因果关系是比较容易处理的线性问题. 但是, 面对复杂系统安全问题属于非线性问题, 很难得到通用解, 一般只能通过数值方法来得到一些特殊解, 而且复杂系统安全问题未必有可行的数学模型描述因果关系. 大数据技术通过寻找相关数据之间的关系, 从而忽略中间过程, 忽略其中的因果细节, 构建认识问题的 “数据模型”, 从宏观上去把握数据之间的相关关系, 使非线性问题有了具体的解决路径. 通过对大数据进行相关性分析来找出事物之间的关联, 既可以避免主观偏见的影响, 又可为研究因果关系打下良好基础. 需指出, 大数据技术重视相关性忽视因果性并不意味着怀疑或否定事物之间的因果关系.

4) 增强系统安全理论模型预测功能: 事故 (尤其是重特大事故) 的发生具有必然性和偶然性, 传统的基于小数据及其线性因果关系的安全理论模型可以解释必然性, 但不能预测偶然性. 面对现代复杂社会技术系统, 系统安全解释和预测都比较复杂. 大数据技术的核心理念是如何利用大数据进行预测, 基于大数据的系统安全理论模型有助于提前、快速的识别与判断系统未来的安全状态, 真正做到系统安全的预测与预控. 此外, 大数据模型来源于海量数据, 安全数据量的增长意味着蕴含的安全信息会更多, 越多的安全信息运用到模型中, 系统安全预测就会越准确.

5) 实现系统安全动态演变仿真建模: 仿真技术在很多领域得到应用, 理论建模是仿真实现的基础. 在系统安全理论建模与仿真领域, 目前还停留在理论建模阶段, 由于理论和技术的限制, 还不能实现系统安全动态演变的仿真模拟. 基于大数据的系统安全理论建模, 一方面可以完善现有安全理论模型在实现仿真方面的不足, 另一方面, 为仿真的实现提供可行的技术路径, 实现系统安全动态演变的仿真建模, 真正实现系统安全的可视化与可感化.

6) 实现安全理论建模的大综合大交叉: 大数据技术的兴起为复杂性科学提供了具体的实现路径, 使得复杂性科学方法论变得可操作. 技术层面从小数据到大数据的变革将从本质上推动科学层面从简单性安全理论建模到复杂性安全理论建模的转变, 基于大数据思维与技术的安全理论建模是安全科学研究从简单性科学研究范式向复杂性科学研究范式转变的重要表现. 大数据技术将打破传统安全数据壁垒和安全思维定式, 极大的促进安全科学共同体之间的交流、汇集与资源共享, 促进各个学科 (如技术科学、自然科学、社会科学、生命科学、系统科学等) 的学者跳出自己所在领域的局限性, 实现安全科学研究的跨学科、跨地区、跨时空的大规模合作, 实现安全数据、安全信息、安全知识和安全科学的资源共享与融合, 推动安全理论建模研究与实践真正走向 “大安全”.

综上分析, 大数据技术的兴起对传统安全科学思维和安全理论建模方法带来了挑战和变革. 在挑战方面, 传统的 “中小系统型” “还原型” “小数据型” “静态型” “解释型” “因果型” “抽象型” “简化型” 和 “ 小综合小交叉

型” 安全理论模型将不能满足愈发复杂的社会技术系统安全需求. 在变革方面, 大数据技术将促使系统安全理论建模面向全体数据、实现还原论与整体性方法的融贯、突出相关关系、增强预测性、实现系统安全动态仿真、实现大综合大交叉. 在这种 “挑战 - 机遇” 机制 (或 “拉 - 推” 机制) 的作用下, 扭转传统安全理论建模缺陷, 促使安全理论建模适应新时代、新技术背景下的复杂系统安全需求.

4　基于大数据的系统安全理论建模范式

为便于论证, 首先给出基于大数据的系统安全理论建模新范式, 见图 3. 从 3 方面对其进行剖析: 基于大数据的安全理论建模原理、技术路径和逻辑主线.

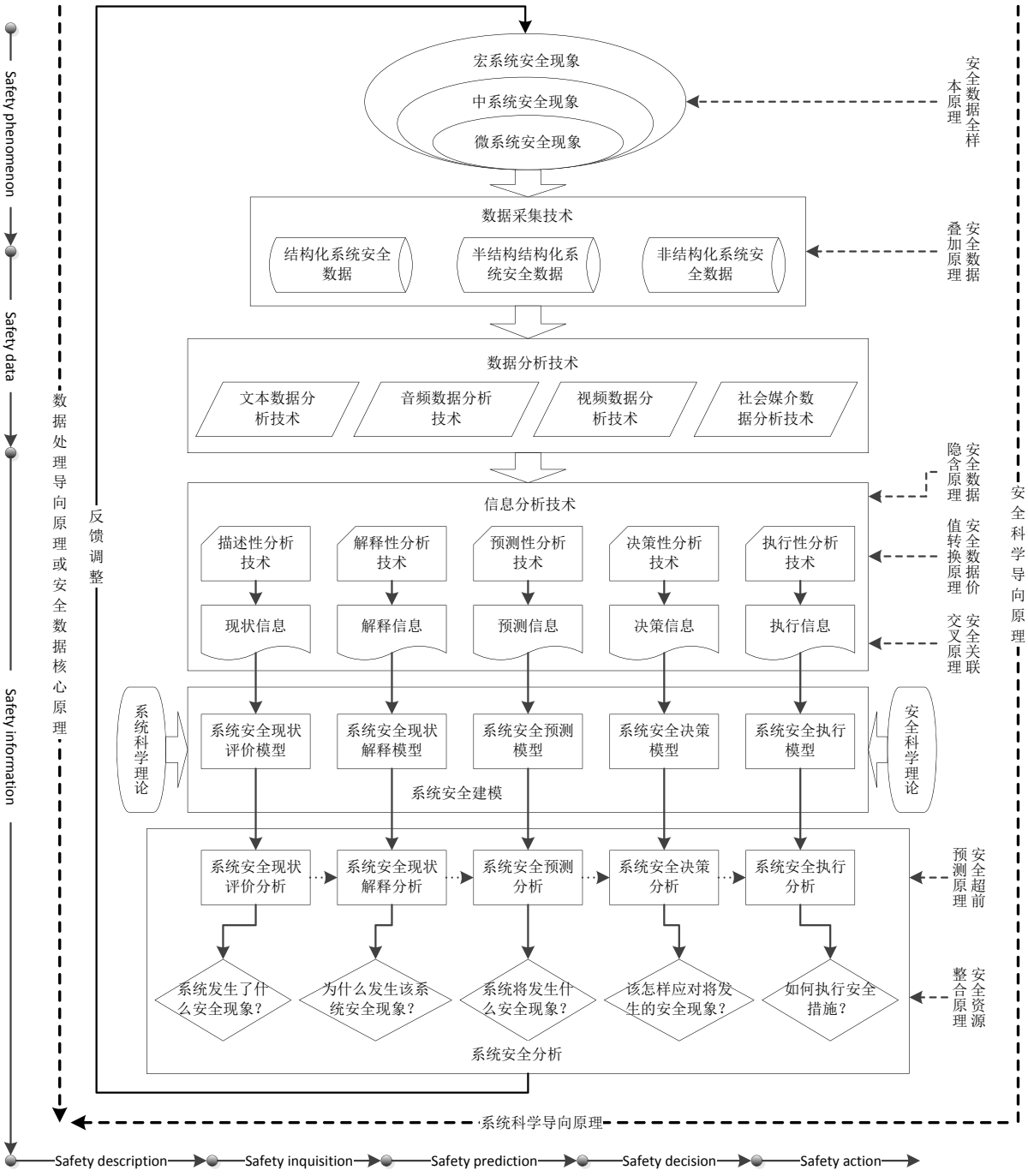

图 3　基于大数据的系统安全理论建模范式

4.1 基于大数据的系统安全理论建模技术路径

大数据分析技术是数据驱动的系统安全理论建模的基础, 基于不同的数据分析技术, 可构建不同的系统安全模型. 根据文献 [28–30] 相关研究, 本文将大数据分析技术分为 “一次分析技术” 和 “二次分析技术”, 其中 “一次分析技术” 直接面对原始数据 (如文本、声音、视频等), 将采集到的结构化数据、半结构化数据和非结构化数据转化为可用信息, 如文本数据分析技术 (text analytics)、音频数据分析技术 (audio analytics)、视频数据分析技术 (video analytics)、社会媒介数据分析技术 (social media analytics) 等. “二次分析技术” 是用于提取有用信息, 按照不同的数据分析目的以及不同的分析过程, 可将 “二次分析技术” 分为: ①描述性分析技术 (descriptive analytics)[31], 主要用于评估系统当前的安全状态; ②调查性分析技术 (inquisitive analytics)[32], 主要用于解释系统当前的安全状态; ③预测性分析技术 (predictive analytics)[33], 主要用于预测系统未来的安全状态; ④决策性分析技术 (prescriptive analytics)[31], 主要用于制定备选方案和优选方案; ⑤执行性分析技术 (pre-emptive analytics)[34], 主要用于确定、采取和指导预防性措施的实施.

此外, 从系统工程视角可将系统安全分析分为: 系统安全现状评价、解释、预测、决策、控制和执行等. 因此, 整合不同的大数据分析方法和不同的系统安全分析目的, 可将基于大数据的系统安全模型分为: ①系统安全现状评价模型; ②系统安全解释模型; ③系统安全预测模型; ④系统安全决策模型; ⑤系统安全执行模型.

4.2 基于大数据的系统安全理论建模逻辑主线

系统安全理论模型或系统安全研究的主要功能可以分为 3 个模块: ①面向过去, 分析导致目前安全现象的原因, 进行系统事故致因调查与分析; ②面向现在, 对目前的系统安全现象做出评价, 进行系统安全管理; ③面向未来, 预测未来的安全现象、并做出决策, 进行系统安全预测、决策与执行. 这 3 个模块都是基于系统目前的安全现象做出的. 更重要的是, 安全数据是安全现象的直接体现. 因此, 以系统安全现象作为系统安全理论建模的逻辑起点具有科学性. 根据不同的建模目的和不同的系统粒度, 系统安全现象又可分为微系统、中系统和宏系统安全现象. 可将上述过程的逻辑主线概括为 “安全现象 (safety phenomenon)→ 安全数据 (safety data)”.

根据数据研究 “DIKW (data-information-knowledge-wisdom)” 框架 [35], 以及文献 [13–16] 对大数据应用于安全科学领域和事故调查的相关研究, 可将这一过程概括为 “安全数据 (safety data)→ 安全信息 (safety information)→ 安全洞见 (safety insight)”, 其中安全洞见 (safety insight) 包括安全知识 (safety knowledge) 和安全智慧 (safety wisdom) 或安全科学 (safety science). “安全数据 → 安全信息” 过程主要是从海量安全数据中提取有用安全信息 (现状信息、解释信息、预测信息、决策信息和执行信息). “安全信息 → 安全洞见” 过程是指从获取的安全信息中提炼安全知识和安全智慧, 作为安全理论建模的基础. 又根据 “3.1 基于大数据的系统安全理论建模技术路径”, 可将系统安全行为归纳为 “安全描述 (safety description)→ 安全解释 (safety inquisition)→ 安全预测 (safety prediction)→ 安全决策 (safety decision)→ 安全执行 (safety action)”. 综上, 可将系统安全建模路径 (或逻辑主线) 可以概括为 “safety phenomenon → safety data → safety information → safety description → safety inquisition → safety prediction → safety decision → safety action”.

4.3 基于大数据的系统安全理论建模原理

大数据应用于系统安全理论建模需要以大数据技术、信息技术、系统科学、安全科学等基础技术与学科为理论基础, 还需要这些基础技术与学科融合而成的具体的建模原理作为指导. 在文献 [14–16] 对大数据应用于安全科学领域原理研究的基础上, 本文加入系统科学导向原理, 将基于大数据的系统安全建模原理归纳为: 安全数据全样本原理、安全数据叠加原理、安全数据核心原理 (或数据处理导向原理)、安全数据隐含原理、安全数据价值转化原理、安全关联交叉原理、安全资源整合原理、安全超前预测原理、安全科学导向原理、系统科学导向原理 (详细论述见文献 [14–16]).

各个原理在建模过程中的应用简述如下 (见图 3): ①安全数据全样本原理, 主要用于数据采集阶段, 根据所研究的系统安全问题和系统粒度, 可将数据分为微观层面安全数据、中观层面安全数据和宏观层面安全数据; ②安全数据叠加原理, 主要用于依据数据特点以及数据处理技术对数据进行初选和分类; ③安全数据

核心原理 (或数据处理导向原理), 即去除 “噪音” 类数据, 以大数据处理一般流程为导向; ④安全数据隐含原理、安全数据价值转化原理和安全关联交叉原理, 主要用于从安全数据提取和挖掘有用安全信息; ⑤安全资源整合原理, 基于大数据的系统安全建模涉及传感器技术、物联网技术、数据传输技术、信息技术、系统工程、安全工程技术等, 需要在建模过程中整合现有资源与理论; ⑥安全超前预测原理, 大数据技术最主要的功能是预测预报, 这也是基于大数据的系统安全模型主要功能; ⑦安全科学导向原理, 基于大数据的系统安全建模是以目地 (系统安全) 导向、以数据为核心的, 因此, 在建模过程中需要安全科学理论指导; ⑧系统科学导向原理, 贯穿于整个系统安全理论建模, 将 “数据” 和 “安全” 相结合, 此外, 安全理论模型本身就是一个系统, 其建模过程属于系统工程, 因此, 理应需要系统科学原理.

4.4　新范式的有效性与科学性检验

为了论证所提出的基于大数据的系统安全理论建模新范式的有效性与科学性, 选择 2 个实例进行实证分析.

实例 1: 基于大数据的行为安全管理模型 [18].

以建筑施工现场为例, 建筑工人每天产生大量行为数据, 但这些数据都属于半结构化数据和非结构化数据. 为了克服传统的行为安全管理模型诸多缺陷, 基于大数据构建和开发了一个行为安全管理模型. 该模型以建筑施工现场人员的操作行为 (微系统层面安全现象) 为逻辑起点, 通过智能摄像技术实时监控施工人员的操作行为 (见图 4 左半部分), 获得非结构化的行为数据 (安全数据), 这些具有语义信息的行为数据储存于 Google 开发的 Hadoop 分布式文件处理系统 (Hadoop distributed file system, HDFS), 并通过该系统的图像分析技术分析和提取操作工的行为特征 (安全信息), 和预先设计好的 “行为风险知识库” 中的相关数据进行比较分析 (描述性分析技术)(见图 4 右半部分), 自动判断所感知到的不安全行为, 并进行警示. 该模型可以实现对不安全行为的实时、可视化、自动化监督与管理, 并可以通过大量数据的积累, 实现对模型的修正与完善, 还可以发现传统行为安全模型不能发现的不安全行为模式 (见图 5). 该模型的详细描述参考文献 [18].

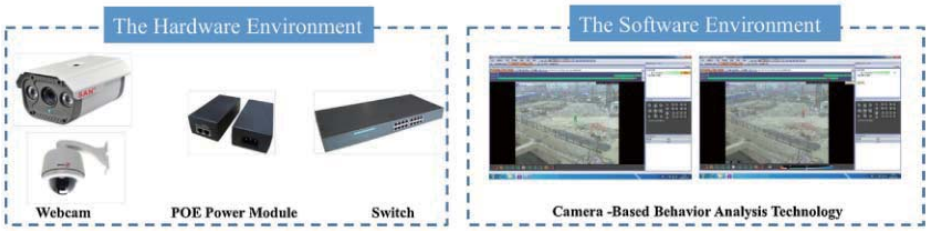

图 4　文献 [18] 构建的基于大数据的行为安全模型软件、硬件展示

图 5　文献 [18] 构建的基于大数据的行为安全模型分析结果展示

实例 2: 基于大数据的城市公路实时交通运行和安全监控与改善模型 [19].

为了充分利用城市智能交通系统产生的大量交通数据, 减少拥堵和碰撞风险, 改善城市高速公路的系统性能, 构建了基于大数据的城市公路实时交通运行和安全监控与改善模型. 通过随机森林 (random forest) 数据挖掘与分析技术和贝叶斯推理分析技术, 对交通拥堵进行实时安全分析, 揭示交通动态情况对碰撞事故的影响, 实现对公路系统交通情况的实时预测预报. 发现城市高速公路拥挤具有高度地局部性和时效性, 正如

预期的那样, 早上和晚上的高峰时段是最拥挤的时间段. 辨识出的间接 (高峰时段, 更高的交通量和更低的速度上游的崩溃地点) 和直接 (较高拥挤指数下游的崩溃地点) 拥堵指标证实了拥堵对后端碰撞可能性的重大影响. 并引入在结构可靠性分析中广泛使用的一阶可靠性方法, 根据拥堵密度确定合适的安全预警触发时间. 有关该模型的详细描述见文献 [19].

通过上述两个实例分析论证了所提出的新范式的可行性与科学性, 新范式可为基于大数据的系统安全理论建模提供理论参考.

5 结论

1) 数据和信息系统在安全生产活动中既是客体又是主体, 具有主客体双重性质. 在进行系统安全理论建模时, 必须考虑数据和信息在安全生产活动中的主体角色. 数字和信息已经成为人机交互的核心纽带, 数字技术和信息技术成为系统安全分析与控制的关键手段, 系统安全分析需要关注多人 (人群)、多机 (机群).

2) 构建由数据维、系统维和安全维组成的系统安全理论建模范式 “数据系统安全” 3 维框架. 将数据驱动的系统安全研究分为 3 范式: “经验型 - 微系统 - 古典安全” 范式、“小数据 - 中系统 - 近代安全” 范式与 “大数据 - 宏系统 - 大安全” 范式, 或 “大数据 - 大系统 - 大安全” 范式.

3) 分析现有 9 类 (中小系统型、解释型、还原型、小数据型、静态型、简化型、因果型、抽象型和小综合小交叉型) 在大数据时代面临的挑战, 从面向全体数据、实现还原论方法与整体性方法的融贯、突出相关关系、增强预测性、实现系统安全动态仿真、实现大综合大交叉 6 个方面分析大数据给系统安全理论建模带来的机遇.

4) 提出基于大数据的系统安全理论建模新范式, 并解析: ①在技术路径方面, 将大数据分析技术分为 “一次分析技术”(文本数据分析技术、音频数据分析技术、视频数据分析技术、社会媒介数据分析技术等) 和 “二次分析技术”(描述性分析技术、调查性分析技术、预测性分析技术、决策性分析技术、执行性分析技术). ②在逻辑主线概括为: 安全现象 → 安全数据 → 安全信息 → 安全描述 → 安全解释 → 安全预测 → 安全决策 → 安全执行. ③将建模原理归纳为: 安全数据全样本原理、安全数据叠加原理、安全数据核心原理 (或数据处理导向原理)、安全数据隐含原理、安全数据价值转化原理、安全关联交叉原理、安全资源整合原理、安全超前预测原理、安全科学导向原理、系统科学导向原理, 并分析各个原理在建模过程中的作用阶段. 最后通过 2 个实例论证所提出的新范式的可行性和科学性.

参考文献

[1] 黄浪, 吴超, 贾楠. 安全理论模型构建的方法论研究 [J]. 中国安全科学学报, 2016, 26(12): 1–6.
Huang L, Wu C, Jia N. Study on methodology of safety theoretical model construction[J]. China Safety Science Journal, 2016, 26(12): 1–6.

[2] 杨冕, 吴超. 安全学演绎逻辑体系的构造 [J]. 系统工程理论与实践, 2016, 36(10): 2712–2720.
Yang M, Wu C. Innovation of the deductive logic system of safety theories[J]. Systems Engineering — Theory & Practice, 2016, 36(10): 2712–2720.

[3] Nature. Big data[EB/OL]. [2017-07-20]. http://www.nature.com/news/specials/bigdata/index. html.

[4] Science. Special online collection: Dealing with data[EB/OL]. [2017-07-20]. http://www.sciencemag.org/site/special/data/.

[5] Hey T. The fourth paradigm — Data-intensive scientific discovery[C]// International Symposium on Information Management in a Changing World. Springer, Berlin, Heidelberg, 2012.

[6] Wolfert S, Lan G, Verdouw C, et al. Big data in smart farming — A review[J]. Agricultural Systems, 2017, 153: 69–80.

[7] Walker G, Strathie A. Big data and ergonomics methods: A new paradigm for tackling strategic transport safety risks[J]. Applied Ergonomics, 2016, 53: 298–311.

[8] Gunasekaran A, Papadopoulos T, Dubey R, et al. Big data and predictive analytics for supply chain and organizational performance[J]. Journal of Business Research, 2016, 70: 308–317.

[9] Wang H, Xu Z, Fujita H, et al. Towards felicitous decision making: An overview on challenges and trends of Big Data[J]. Information Sciences, 2016, 367–368: 747–765.

[10] Chang R M, Kauffman R J, Kwon Y O. Understanding the paradigm shift to computational social science in the presence of big data[J]. Decision Support Systems, 2014, 63(3): 67–80.

[11] Guo H, Wang L, Liang D. Big earth data from space: A new engine for Earth science[J]. Science Bulletin, 2016, 61(7): 505–513.
[12] 郭旦怀, 催文娟, 郭云昌, 等. 基于大数据的食源性疾病事件探测与风险评估 [J]. 系统工程理论与实践, 2015, 35(10): 2523–2530.
Guo D H, Cui W J, Guo Y C, et al. Foodborne disease event detection and risk assessment based on big data[J]. Systems Engineering — Theory & Practice, 2015, 35(10): 2523–2530.
[13] Huang L, Wu C, Wang B, et al. A new paradigm for accident investigation and analysis in the era of Big Data[J]. Process Safety Progress, http://dx.doi.org/10.1002/prs.11898.
[14] Ouyang Q M, Wu C, Huang L. Methodologies, principles and prospects of applying big data in safety science research[J]. Safety Science, 2018, 101: 60–71.
[15] 欧阳秋梅, 吴超, 黄浪. 大数据应用于安全科学领域的基础原理研究 [J]. 中国安全科学学报, 2016, 26(11): 13–18.
Ouyang Q M, Wu C, Huang L. Research on basic principles of applications of big data in field of safety science[J]. China Safety Science Journal, 2016, 26(11): 13–18.
[16] 王秉, 吴超. 基于安全大数据的安全科学创新发展探讨 [J]. 科技管理研究, 2017, 37(1): 37–43.
Wang B, Wu C. Study on the innovation research of safety science based on the safety big data[J]. Science and Technology Management Research, 2017, 37(1): 37–43.
[17] Shu Y, Ming L, Cheng F, et al. Abnormal situation management: Challenges and opportunities in the big data era[J]. Computers & Chemical Engineering, 2016, 91: 104–113.
[18] Guo S Y, Ding L Y, Luo H B, et al. A Big-Data-based platform of workers' behavior: Observations from the field[J]. Accident Analysis & Prevention, 2016, 93: 299–309.
[19] Shi Q, Abdel-Aty M. Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways[J]. Transportation Research Part C: Emerging Technologies, 2015, 58: 380–394.
[20] Gil D, Song I Y, Aldana J F, et al. Big Data. New approaches of modelling and management[J]. Computer Standards & Interfaces, 2017, 54: 61–63.
[21] 宋守信, 陈明利. 关于信息社会安全理论发展的几点思考 —— 甬温线动车事故的启示 [J]. 中国安全科学学报, 2013, 23(3): 140–144.
Song S X, Chen M L. New perspective on safety theory development in information society: Inspiration from Yongwen line train collision in China[J]. China Safety Science Journal, 2013, 23(3): 140–144.
[22] 刘潜. 安全科学和学科的创立与实践 [M]. 北京: 化学工业出版社, 2010.
Liu Q. The establishment and practice of safety science and discipline[M]. Beijing: Chemical Industry Press, 2010.
[23] Stoop J, Kroes J D, Hale A. Safety science, a founding fathers' retrospection[J]. Safety Science, 2017, 94: 103–115.
[24] 黄浪, 吴超. 事故致因模型体系及建模一般方法与发展趋势 [J]. 中国安全生产科学技术, 2017, 13(2): 10–16.
Huang L, Wu C. Structure system of accident-causing model and its modeling methodology and development trend[J]. Journal of Safety Science and Technology, 2017, 13(2): 10–16.
[25] Harvey C, Stanton N A. Safety in System-of-Systems: Ten key challenges[J]. Safety Science, 2014, 70: 358–366.
[26] Qureshi Z H. A review of accident modelling approaches for complex socio-technical systems[C]// Twelfth Australian Workshop on Safety Critical Systems and Software and Safety-Related Programmable Systems. Australian Computer Society, Inc, 2007: 47–59.
[27] 欧阳秋梅, 吴超. 大数据与传统安全统计数据的比较及其应用展望 [J]. 中国安全科学学报, 2016, 26(3): 1–7.
Ouyang Q M, Wu C. On comparison between big data and traditional safety statistics and big data's application prospects[J]. China Safety Science Journal, 2016, 26(3): 1–7.
[28] Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics[J]. International Journal of Information Management, 2015, 35(2): 137–144.
[29] Sivarajah U, Kamal M M, Irani Z, et al. Critical analysis of Big Data challenges and analytical methods[J]. Journal of Business Research, 2016, 70: 263–286.
[30] 杨青, 武高宁, 王丽珍. 大数据: 数据驱动下的工程项目管理新视角 [J]. 系统工程理论与实践, 2017, 37(3): 710–719.
Yang Q, Wu G N, Wang L Z. Big data: A new perspective of the engineering project management driven by data[J]. Systems Engineering — Theory & Practice, 2017, 37(3): 710–719.
[31] Joseph R C, Johnson N A. Big data and transformational government[J]. IT Professional, 2013, 15(6): 43–48.
[32] Bihani P, Patil S. A comparative study of data analysis techniques[J]. International Journal of Emerging Trends & Technology in Computer Science, 2014, 3(2): 95–101.
[33] Waller M A, Fawcett S E. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management[J]. Journal of Business Logistics, 2013, 34(2): 77–84.
[34] Szongott C, Henne B, Voigt G. Big data privacy issues in public social media[C]// 6th IEEE International Conference on Digital Ecosystems Technologies (DEST), 2012: 1–6.
[35] Baskarada S, Koronios A. Data, information, knowledge, wisdom (DIKW): A semiotic theoretical and empirical exploration of the hierarchy and its quality dimension[J]. Australasian Journal of Information Systems, 2013, 18(1): 2109–2112.