

## 基于数据挖掘的化工生产事故致因主题抽取

牛毅 樊运晓 高远

(中国地质大学(北京)工程技术学院 北京 100083)

**摘要:** 为充分挖掘化工生产事故数据中的有效信息和潜在规律,提高对化工事故认知水平,针对某化工集团2010—2016年共1 578起事故数据,利用社会网络分析等方法揭示事故要素间的关联关系;运用潜在狄利克雷分配(LDA)模型进行事故聚类,并抽取到5个事故致因主题。研究结果表明:LDA主题模型等数据挖掘技术能有效挖掘大量事故数据中的潜在信息;5个事故致因主题中4个涉及到人因或组织层面的缺陷;员工注意力不集中和现场风险管理不足这2个致因主题间具有较强相关性;员工注意力不集中、现场风险管理不足以及设备问题是导致事故发生的主要原因。

**关键词:** 化工事故;文本数据;数据挖掘;潜在狄利克雷分配(LDA);事故致因

**中图分类号:** X928.0 **文献标志码:** A **文章编号:** 1673-193X(2019)-10-0165-06

### Topic extraction on causes of chemical production accidents based on data mining

NIU Yi, FAN Yunxiao, GAO Yuan

(School of Engineering & Technology, China University of Geosciences Beijing, Beijing 100083, China)

**Abstract:** In order to fully exploit the effective information and potential laws in the data of chemical production accidents, and improve the cognitive level of chemical accidents according to the data of 1 578 accidents in a chemical industry group from 2010 to 2016, the association relationship between the accident elements was revealed by means of social network analysis. The Latent Dirichlet Allocation (LDA) model was applied to conduct the accident clustering, and five topics of accident causes were extracted. The results showed that the data mining techniques such as LDA topic model could effectively mine the potential information in a large amount of accident data. Among the five topics of accident causes, four topics involved the human factors or organizational defects. Two topics of accident causes including the lack of concentration of employees and the inadequate on-site risk management had a strong correlation. The lack of concentration of employees, the inadequate on-site risk management and the equipment problems were the main causes of accidents.

**Key words:** chemical accidents; text data; data mining; Latent Dirichlet Allocation (LDA); accident cause

### 0 引言

近年来,尽管我国化工安全生产形势整体趋好,但化工安全生产事故依然多发<sup>[1]</sup>。从历史事故中学习经验对进一步遏制事故的发生具有重要意义,然而人们往往更关注重大事故,而忽视小事故。海因里希曾在统计大量事故后提出著名的事故金字塔理论:伤亡、轻伤和无伤害事故的比例为1:29:300<sup>[2]</sup>,即1起严重事故背后必然有更多的小事故发生,要防止严重事故发生必须先减少或消除小事故,小事故同样具有重要的研究价值<sup>[3]</sup>。在化工生产过程中,企业积累了大量的小型事故

文本数据,但受数据量大、记录粗糙以及相应信息处理能力不足等因素限制,数据价值未得到充分利用,进而制约了对事故的深入研究。因此,如何从这些杂乱的事故信息中高效挖掘潜在价值,已成为亟待解决的问题。

当前针对化工事故数据应用的研究,还主要停留在对事故类型、事发区域或时段的简单统计分析方面<sup>[4]</sup>。面对复杂事故的发生机理,单一的统计指标不能充分描述事故的特征<sup>[5]</sup>。在此背景下,一些学者尝试通过数据挖掘技术对事故数据进行事故分类<sup>[6]</sup>、事故聚类<sup>[5]</sup>和事故潜在关联规则挖掘<sup>[7]</sup>等方面研究,然而这些研究大多集中于道路交通<sup>[8]</sup>和建筑施工<sup>[9]</sup>等领域,在化工安全生

收稿日期: 2019-05-10; 数字出版日期: 2019-10-26

\* 基金项目: 国家自然科学基金项目(51474193)

作者简介: 牛毅,硕士研究生,主要研究方向为工业安全管理、事故致因分析。

通信作者: 樊运晓,博士,教授,主要研究方向为事故预防、系统安全工程、风险管理。

产事故中的应用还相对较少。相较于数值型数据,文本数据作为事故记录的主要形式,蕴含着事故更深入的细节,但由于结构复杂、记录粗糙等原因,对大量化工事故文本数据系统的挖掘则更为缺乏。

基于上述问题,本文引入可从大量事故文本数据中提取主题和关键信息的数据挖掘方法:潜在狄利克雷分配(Latent Dirichlet Allocation, LDA)主题模型和社会网络分析法,以某化工企业 1 578 起小型事故为样本,利用该方法进一步挖掘事故致因及其分布规律,实现对化工事故的更深入分析,为企业事故预防工作提供有力支持。

## 1 数据挖掘

数据挖掘是指通过统计、机器学习等方法,从大量数据中挖掘对研究或决策具有深层价值信息的过程。

### 1.1 化工安全生产事故数据特征

目前,文本数据是各行业安全生产事故的最主要记录形式之一,其中重特大事故会以事故调查报告的形式详细剖析,然而数量上占绝大多数的轻伤或无人员伤亡的小事故往往只进行简单的记录,由于其记录粗糙、数量庞大以及不被重视等原因,针对这些文本数据的挖掘还十分缺乏。以化工行业为例,化工安全生产事故文本数据通常是以描述性的文字对事故发生过程及原因的记录,记录虽然粗糙,但蕴含着事故经过和生产一线工作状况等关键细节,同样具有巨大的利用价值和挖掘空间。

### 1.2 LDA 主题模型

潜在狄利克雷分配(Latent Dirichlet Allocation, LDA)是由 Blei 提出的 1 种主题挖掘模型<sup>[10]</sup>,该模型基于一定规则从海量文本中抽取主题,并根据主题分布对文本数据进行聚类,其中主题是指文本数据所蕴含的潜在主旨或核心思想。目前,该模型已广泛应用于图书情报、热点评论主题抽取<sup>[11]</sup>等领域。LDA 也称为三层贝叶斯概率模型,如图 1 所示,其结构由文本、主题和词项 3 部分组成,其中每个文本由一系列服从概率分布的主题构成,主题又由一系列服从概率分布的词项构成<sup>[10]</sup>。

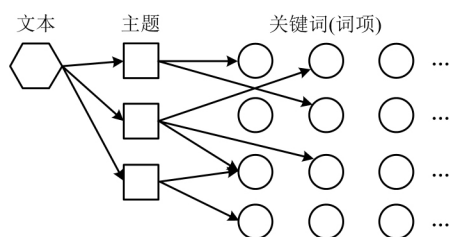


图 1 三层贝叶斯概率结构模型

Fig. 1 Three-layer Bayesian probabilistic structure model

1) 主题生成过程: LDA 主题模型如图 2 所示,图 2

中主要符号含义:  $\alpha$   $\beta$  为狄利克雷先验参数;  $m$  表示第  $m$  个文本;  $\theta_m$  表示文本  $m$  的主题分布;  $n$  表示文本  $m$  的词项数;  $z_{m,n}$  表示文本  $m$  的第  $n$  个词项对应的主题;  $w_{m,n}$  表示文本  $m$  的第  $n$  个词项;  $\varphi_{z_{m,n}}$  表示主题的词项分布;  $K$  表示主题总数;  $M$  表示文本总数;  $k$  表示第  $k$  个主题;  $N_m$  表示第  $m$  篇文档的词项总数;  $Dir$  表示狄利克雷分布;  $Mult$  表示多项式分布。

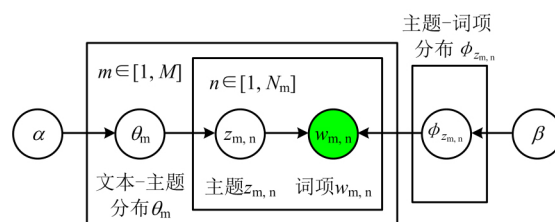


图 2 LDA 主题模型

Fig. 2 LDA probability model

主题详细生成过程为<sup>[12]</sup>:

- ①从狄利克雷分布  $\alpha$  中取样生成文档  $m \in (1, M)$  的主题分布  $\theta_m$ ,  $\theta_m | \alpha \sim Dir(\alpha)$ ;
- ②针对文档  $m$  中的每个词项  $w_{m,n}$ , 其中,  $n \in (1, N_m)$ : a. 从多项式分布  $\theta_m$  中取样生成该词项  $w_{m,n}$  的主题  $z_{m,n} | \theta_m \sim Mult(\theta_m)$ ; b. 从词项分布  $\varphi_{z_{m,n}}$  中取样最终生成词项  $w_{m,n} | \varphi_{z_{m,n}} \sim Mult(\varphi_{z_{m,n}})$ , 其中  $\varphi_{z_{m,n}} | \beta \sim Dir(\beta)$ 。在给定参数  $\alpha$   $\beta$  情况下,模型的联合概率分布为:

$$p(\theta, \varphi, z, w | \alpha, \beta) = \left( \prod_{k=1}^K p(\varphi_k | \beta) \right) p(\theta | \alpha) \prod_{n=1}^N p(z_n | \theta) p(w_n | z_n, \varphi_{z_n}) \quad (1)$$

2) 参数估计: 图 2 中, 阴影圆圈表示观测变量, 非阴影圆圈表示潜在变量, 模型求解前需要对潜在变量进行估计。本文中, 模型的参数估计主要基于变分推断 EM 算法。其原理为: 通过近似隐藏变量的后验分布来简化问题, 通过寻求最大似然解来进行参数估计<sup>[12]</sup>。

### 1.3 Tf-idf 算法

事故记录中每个关键词是蕴含事故信息的基本单元, 考虑到一些词频少的关键词可能携带更多事故信息, 因此利用 Tf-idf (Term frequency-inverse document frequency) 算法<sup>[13]</sup>对事故关键词进行加权处理。算法原理为: 一个关键词在整个数据集中出现频率不高, 但在某起事故中多次出现, 则认为该关键词既对该起事故有很好的代表性, 又对其他事故有很好的区分能力, 那么此关键词在该事故中具有较高的权重。通过加权处理可有效避免这些关键事故信息的流失。

### 1.4 社会网络分析法

社会网络分析法旨在对系统要素间的关系进行量

化研究<sup>[14]</sup>, 一个网络通常由“点”和“边”构成, “点”表示各个要素, “边”表示要素间的关系, “边”的粗细程度表示关联程度。事故致因要素间并非简单的线性关系, 而是构成复杂的关系网络, 为进一步发掘事故要素间的关系, 基于社会网络分析法, 将关键词看作“点”, 关键词之间的关联看作“边”, 构建事故关键词网络。算法原理为: 在整个事故文本数据集中, 若 2 个关键词同时出现

在一个句子的频率超过一定值时, 则认为这 2 个关键词有关联, 该频率大小反映了关联度的强弱。

## 2 化工事故致因挖掘流程

化工事故致因挖掘主要包括数据收集与处理、事故致因挖掘、结果可视化分析等 3 个主要环节, 具体流程如图 3 所示。

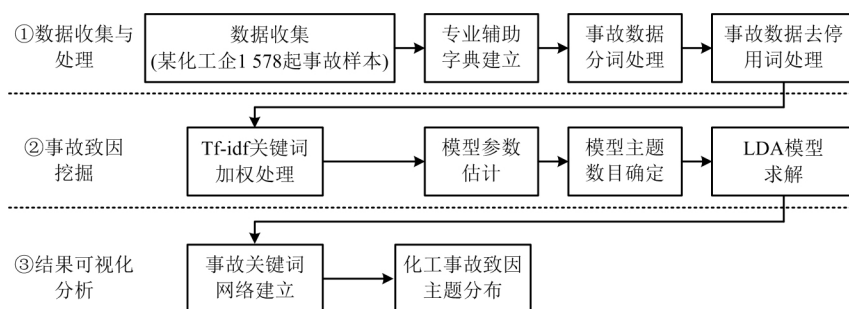


图 3 化工事故主题抽取流程

Fig. 3 Chemical accident theme extraction process

### 2.1 数据收集与处理

本研究的数据来源为国内某大型化工集团 2010—2016 年期间的 1 578 起事故记录, 该企业为中国最具代表性的化工企业之一。企业拥有较完善的事故信息采集系统, 1 578 起事故中大多为轻伤和设备损坏事故(少量死亡、重伤事故), 记录内容为事故发生后第一时间对事故经过、生产环境和事故原因的文字性描述, 蕴含着事故发生时的关键细节, 且所有数据的记录形式统一, 利于对该数据集进行系统地潜在规律挖掘。原始事故数据示例见表 1。

表 1 原始事故数据(示例)

Table 1 Accident data after modification( example)

文本文档	事故经过及原因描述
2010108.txt	2010 年 8 月 4 日上午 9 时, 研发部门研发员准备进行改性塑料试验, 在清理切粒机滚刀时, 用手盘动刀片, 不注意将其右环指轧伤, 伤后疼痛, 流血, 通附院诊断为右环指末节指骨骨折。促成因素: 清理切粒机时, 用手盘动切粒机刀片; 用手清理。根本原因: 操作时注意力不集中, 未意识到有伤害的可能性; 管理上存在缺陷

为达到模型计算的要求, 同时提高模型效率及准确性, 事故致因挖掘前, 需将原始事故数据进行分词、去停用词等一系列处理:

1) 分词处理: Jieba 分词包是 1 个基于 python 实现的分词工具, 借助该工具对事故文本数据进行分词处理。由于要分析数据的背景为化工事故, 为了避免分词过程中无法识别某些化工或安全专业名词的情况, 结合

化工企业相关规范, 制定了适用于化工安全事故分析的辅助字典, 见表 2。

表 2 化工安全专业领域字典(部分)

Table 2 Chemical safety professional domain dictionary (partial)

词汇	
化工领域	开车 试车 停车 物料 管线 炼油厂 热油 聚合 分馏 净化 防喷器 换热器……
安全领域	安全 风险 警惕性 意识 注意力 危害 应急预案 故障 违章作业 隐患……

2) 去停用词: 事故记录中存在许多对事故分析无意义的词项, 为提高模型效率, 将这些词项加入停用词表, 并将其过滤, 即去停用词处理。

### 2.2 事故致因挖掘

通过对原始数据处理, 得到适合模型计算的 1 578 起事故数据。基于 Python 编程语言, 利用前文所介绍的 LDA 主题模型、Tf-idf 算法以及社会网络分析法, 对 1 578 起事故数据进行事故致因主题抽取、关键词加权以及建立关键词关系网络等工作。其具体流程如图 3 中②部分所示。

### 2.3 结果可视化分析

为帮助企业管理者更直观地了解企业安全状况及事故致因分布规律, 将事故关键词关联规则、事故致因主题分布等抽象结果以文字—图像交互的形式展现出来, 利用 Gephi 等可视化工具, 绘制事故关键词关系网络、事故致因主题分布图等可视化结果。

### 3 结果及分析

#### 3.1 事故关键词权重矩阵

事故记录中每个关键词所蕴含信息的重要程度不同,因此利用 Tf-idf 算法,建立文档—关键词权重矩阵,见表3。矩阵中的权重代表了关键词在事故中的重

要程度,权重越高重要度越高。例如,事故‘3’中,关键词“不慎”的权重较高,说明注意力问题可能是导致该事故发生的主要原因。

#### 3.2 事故关键词关系网络

基于社会网络分析法,利用 Gephi 软件绘制关键词关系网络图,如图4所示。整个事故集中,重要度越高

表3 关键词权重矩阵(部分)  
Table 3 Keyword weight matrix (partial)

事故编号	关键词							
	不慎	人员	休息	作业	停车	更换	检查	...
1	0	0.360 968	0	0.409 856	0	0	0	
2	0	0	0	0	0	0	0.429 994	
3	0.811 708	0	0	0	0	0	0	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...
1 577	0	0	0	0	0.140 666	0.173 619	0.210 814	
1 578	0.307 440	0.247 843	0	0.281 410	0	0	0.237 170	

的关键词,越靠近关系网络的中心位置,且关键词标签越大。相关的关键词由直线相连,直线的粗细则代表了关联程度的大小。由图4可知,“操作”、“检查”、“设备”等关键词位于整个关系网络的枢纽,尤其是“操作”,位于关系网络的最中心,说明不安全操作仍是事故预防工作中需重点防范的对象。

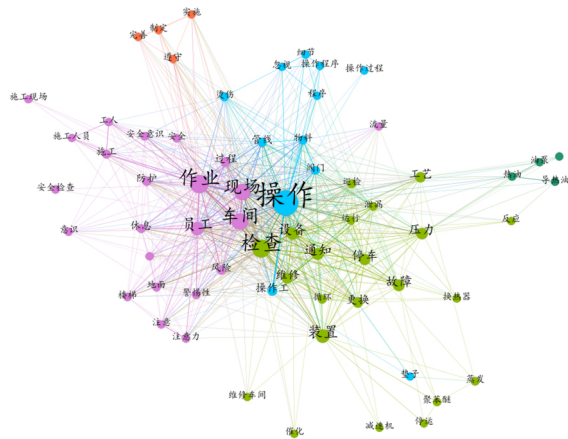


图4 关键词关系网络

Fig.4 Keywords Relational network

#### 3.3 事故致因主题结果分析

##### 1) 主题数确定

本文采用困惑度指标<sup>[15]</sup>来确定主题数目,困惑度指标越小说明建模能力越好,则主题数目为最优,计算公式为:

$$perplexity = \exp \left\{ - \frac{\sum_m \log(P(w_m))}{\sum_m N_m} \right\} \quad (2)$$

式中:  $perplexity$  表示困惑度;  $m$  表示第  $m$  篇文档;  $P(w_m)$  表示第  $m$  篇文档每个单词的概率;  $N_m$  表示第  $m$  篇文档的词项总数。困惑度随主题数目变化时的曲线如图5所示,从图5中可以看出,当主题数为5时,困惑度指标曲线处于低谷,此时模型效果最优,因此将主题数确定为5个。

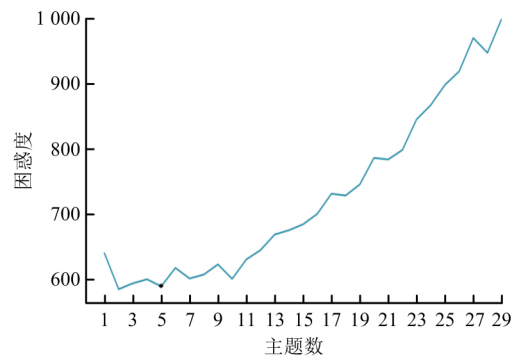


图5 困惑度指标

Fig.5 Perplexityindicator

##### 2) 事故致因主题抽取结果

事故致因主题抽取结果见表4,通过主题模型求解,得到事故致因主题和关键词,并分析每个主题关键词之间的关系得到事故致因总结。其中,事故致因主题揭示了造成事故发生的主要原因类别,主题关键词反映了导致事故发生的关键细节,事故致因总结则是根据关键词对主题进一步细分为若干致因。

运用可视化工具 pyLDAvis 绘制事故致因主题分布图,如图6所示。图6中左侧,圆圈代表不同的主题,主

题标号与表 4 中相对应,圆圈大小代表了每个主题包含的事故数量,圆圈之间的距离代表了主题之间的关联程度。图右侧为频率最高的 30 个关键词列表。由图 6 可知,主题①数量最多,主题①、②距离接近且有重合部分,说明相关性较强,主题④、⑤则数量最少。对各事故致因主题具体分析如下:

表 4 主题抽取结果  
Table 4 Topic extraction results

事故致因主题	主题关键词	事故致因总结
①员工注意力主题	员工 检查 工作 警惕性 警惕 车间 楼梯 不慎 骨折 地面 注意 注意力 休息 操作 操作过程	①员工警惕性不强;②注意力不集中;③休息不足; ④楼梯、车间等区域危险状态
②作业现场风险管理主题	作业 员工 车间 工作现场 操作 风险 过程 维修 管线 安全意识 检查 防护 烫伤	①作业现场安全检查不到位;②安全防护不到位; ③维修不及时;④未识别风险
③设备主题	设备 检查 维修故障 更换 巡检停车现场 通知 运行操作操作工 物料 泄漏 阀门	①设备检查与维修;②设备故障;③变更管理
④制度主题	制度 忽视 操作 细节 程序 操作程序 班组 遵守 施工 施工人员 实施 制定 完善 工人	①制度不完善;②制度制定与实施;③施工人员对 制度不遵守;④对操作程序等细节忽视
⑤化学物质及工艺主题	装置 热油 流量 反应 停车 垫子 聚苯醚 苯醚 换热 蒸发 工艺 导热油 压力 换热器 油泵 催化 故障 停运	①化学物质性质;②化学工艺过程中换热、蒸发、催 化反应;③化学装置故障

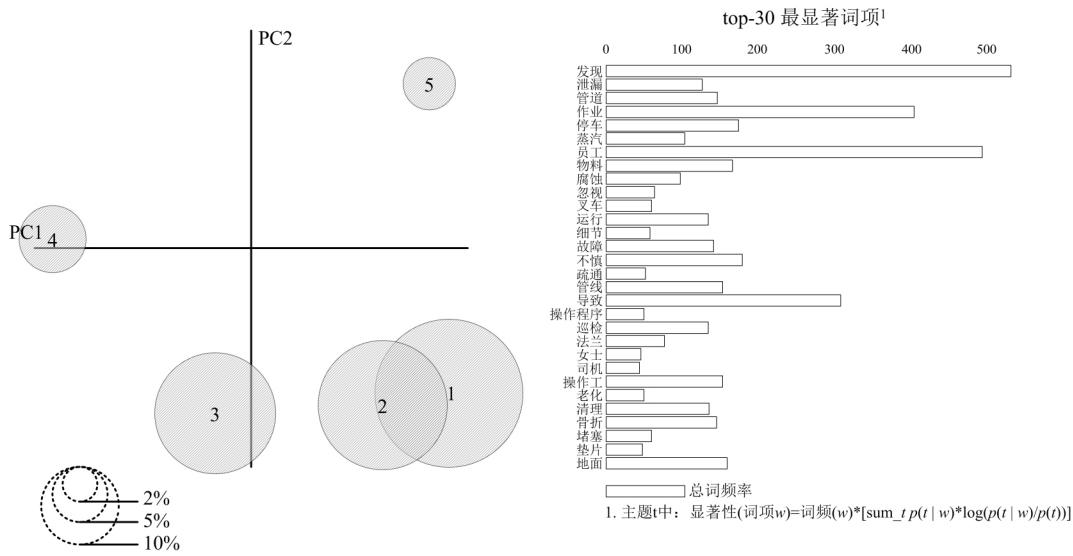


图 6 主题分布可视化  
Fig. 6 Theme distribution visualization

1) 员工注意力主题。由表 4 可知,在员工注意力主题方面,导致事故发生的主要表现为:在车间、楼梯等危险性较高区域,由于员工警惕性不强或注意力不集中而导致了事故发生。关键词“休息”反映出,休息不足或疲劳是造成注意力不集中的重要原因,“车间”“楼梯”则是员工在注意力不集中时事故高发的区域。在图 4 关键词网络中,该主题关键词处于中心位置,说明在关系网络中具有较高重要度。由图 6 可以看出,主题①数量最多,且与主题②有较高相关性。因此,从数量角度看,员工注意力不集中是造成事故发生的重要原因;从相关性角度看,员工注意力问题与作业现场风险管理问题之

间存在较强关联。

2) 作业现场风险管理主题。作业现场安全检查不到位、安全防护缺失、维修不及时以及未识别风险是作业现场风险管理主题的主要表现。由图 4 和图 6 可知,该主题在分布和数量上都与主题①较为相似,说明作业现场风险管理不足也是造成事故发生的重要原因之一,且与主题①具有一定关联性。因此,可以启发管理者在事故预防工作中,关注并切断 2 致因主题之间的某些关联要素,以提高预防工作的效果。例如,通过完善作业现场的风险管理,降低“车间”、“工作现场”的不安全状态,进而提高生产系统的容错率。

3) 设备主题。设备主题由设备检查与维修、设备故障和变更管理等问题组成。通过对主题关键词分析,“检查”“维修”“巡检”等关键词突出的是设备维护不足问题,这就导致设备的不安全状态得不到及时发现和消除,从而产生安全隐患。“更换”等关键词则说明企业在变更管理工作中存在缺陷,未能消除由于设备变更而引起的事故隐患。由图 6 可知,该主题数量较多,也是导致事故发生的主要原因之一。

4) 制度主题。从制度主题的关键词可以看出,主题主要由制度不完善、制度制定与实施、施工人员对制度不遵守等问题组成。从主题分布图中可知,相较于前 3 个主题,该主题数量明显较少,且与其他主题的相关性较弱。但制度本身是较为抽象的概念,往往在事故记录中难以直接体现,尽管如此,制度主题依然是 5 个事故致因主题之一,足以说明安全制度或管理体系问题的严峻性和重要性。

5) 化学物质及工艺主题。上述 4 个致因主题都涉及到人因或组织缺陷,主题⑤则更加侧重于化学物质或工艺过程中本身的性质和反应。从主题关键词推断,“导热油”“聚苯醚”等化学物质和“换热”“蒸发”“催化反应”等工艺过程中释放的能量是造成人员伤亡及装置破坏的主要致因。因此,企业应在适当区域设立防护设施以减轻该类事故对工人及设备造成的危害。

通过聚类结果可以看出,5 个主题中 4 个涉及到人因或组织层面的缺陷,并占据事故致因绝大多数比重,说明事故预防的重点应放在改善人员安全素养和安全管理体系建设等方面。

## 4 结论

1) 通过 LDA 主题模型、社会网络分析法等数据挖掘技术,以某化工企业 2010—2016 年文本类型事故数据为样本,通过挖掘关键词权重、事故构成要素关联关系、事故致因主题等潜在信息,实现了对该类事故数据的进一步识别与认识。

2) 通过 LDA 主题模型抽取到的 5 个主题分别为员工注意力主题、作业现场风险管理主题、设备主题、制度主题、化学物质及工艺主题。分析聚类结果可知,员工注意力不集中、现场风险管理不足、设备检修不及时是导致事故发生的最主要原因,且员工注意力和现场风险管理之间存在较强关联性;人因和组织结构缺陷是导致大量事故发生的根本原因。

3) 本文数据来源为单一化工企业,研究结果对化工行业具有一定参考价值,同时为其他行业事故文本数据挖掘提供了思路,但更系统的事故致因挖掘分析需要更全面的数据支撑。

## 参考文献

[1] 戚萌,朱常龙. 2013—2017 年中国石化及化工行业安全生产现

状及展望[J]. 现代化工, 2019, 39(2): 1-8.

QI Meng, ZHU Changlong. 2013—2017 Sinopec and chemical industry safety production status and prospects[J]. Modern Chemical Industry, 2019, 39(2): 1-8.

[2] HOLLNAGEL E. Safety-I and safety-II: the past and future of safety management [M]. Farnham: Ashgate, 2014: 35-36.

[3] BELLAMY LINDA J. Exploring the relationship between major hazard, fatal and non-fatal accidents through outcomes and causes[J]. Safety Science, 2015(71): 93-103.

[4] 李健,白晓昀,任正中,等. 2011~2013 年我国危险化学品事故统计分析及对策研究[J]. 中国安全生产科学技术, 2014(6): 142-147.

LI Jian, BAI Xiaojun, REN Zhengzhong, et al. Statistical analysis and countermeasures of dangerous chemicals accidents in China from 2011 to 2013 [J]. Journal of Safety Science and Technology, 2014(6): 142-147.

[5] LORENZO, MICAELA, GABRIELE. A combined approach for the analysis of large occupational accident databases to support accident-prevention decision making [J]. Safety Science, 2018(106): 191-202.

[6] RIVAST PAZ M, MARTINJE, et al. Explaining and predicting workplace accidents using data-mining techniques [J]. Reliability Engineering and System Safety, 2011(96): 739-747.

[7] ZHANG Z, HE Q, GAO J, et al. A deep learning approach for detecting traffic accidents from social media data [J]. Transportation Research Part C, 2018(86): 580-596.

[8] ALIKHANI M, NEDAIE A, AHMADVAND A. Presentation of clustering-classification heuristic method for improvement accuracy in classification of severity of road accidents in Iran [J]. Safety Science, 2013(60): 142-150.

[9] YANG MIANG GOH, UBEYNARAYANACU. Construction accident narrative classification: An evaluation of text mining techniques [J]. Accident Analysis and Prevention, 2017(108): 122-130.

[10] BLEI D M, NG AY, JORDAN M J. Latent dirichlet allocation [J]. Journal of Machine Learning Research, 2003(3): 993-1022.

[11] 李静. 基于 LDA 的微博灾害信息聚合——以台风为例 [D]. 武汉: 武汉大学, 2018.

[12] RAMESH N, WILLIAM C, JOHN L. Parallelized variational EM for Latent Dirichlet Allocation: an experimental evaluation of speed and scalability [J]. IEEE International Conference on Data Mining, 2007: 349-354.

[13] 唐明, 朱磊, 邹显春. 基于 Word2Vec 的一种文档向量表示 [J]. 计算机科学, 2016, 43(6): 214-217.

TANG Ming, ZHU Lei, ZOU Xianchun. A document vector representation based on Word2Vec [J]. Computer Science, 2016, 43(6): 214-217.

[14] 朱庆华, 李亮. 社会网络分析法及其在情报学中的应用 [J]. 情报理论与实践, 2008, 31(2): 179-183.

ZHU Qinghua, LI Liang. Social network analysis and its application in information science [J]. Information Theory and Practice, 2008, 31(2): 179-183.

[15] AU ASUNCION, P SMYTH, M WELLING. Asynchronous distributed learning of topic models [J]. Conference on Neural Information Processing Systems, 2011: 81-88.

(责任编辑: 郭利)