

应急预案数字化智能生成的专家系统研究

郝睿轩¹ 黄文才² 陈位泰³

(石河子大学 新疆维吾尔自治区石河子市 832003)

摘要: 新世纪以来,数字化、人工智能、互联网+与各行各业结合并产生深刻影响,在应急管理领域,人工智能的机器学习、数据挖掘等技术的应用已经屡见不鲜,新时代由“人治”向“人与智”的变革,对应急治理体系能力提升有着重大意义,本文笔者根据人工智能产生以来各项技术算法与应急预案数字化研究的理论结合,对现阶段知识库技术、案例推理技术进行总结得出应急预案技术,结合基于 JAVA、自然语言处理与深度学习构建决策支持的本体模型现阶段研究,提出基于知识库+案例推理+智能生成+决策支持的专家系统,对应急预案数字化系统实现智能生成尽可能进行理论验证。

关键词: 应急预案;数字化;专家系统;人工智能

Research on expert system for digital intelligent generation of emergency plan

Hao Ruixuan¹, Huang Wencai², Chen Weitai³

Shihezi University Shihezi City, Xinjiang Uygur Autonomous Region 832003

Abstract: Since the new century, digitalization, artificial intelligence, Internet plus and all walks of life have combined and had a profound impact. In the field of emergency management, the application of artificial intelligence machine learning, data mining and other technologies has become commonplace. In the new era, the transformation from "rule by man" to "people and intelligence" is of great significance to the improvement of emergency management system capacity, In this paper, according to the theoretical combination of various technical algorithms since the emergence of AI and the research on digitalization of emergency plans, the author summarizes the current knowledge base technology and case reasoning technology and obtains the emergency plan technology. Combining the current research on ontology model of decision support based on JAVA, natural language processing and deep learning, the author proposes an expert system based on knowledge base+case reasoning+intelligent generation+decision support, The realization of intelligent generation of digital emergency plan system shall be theoretically verified as far as possible.

Keywords: emergency plan; digitization; Expert system; artificial intelligence

一、研究背景及意义

全球自上世纪九十年代以来,灾害发生数量呈指数型增长,一系列的重大灾害如:美国“911”恐怖袭击、伦敦地铁连环爆炸、SARS和禽流感、印度洋海啸、中国的雪灾更是对世界各国以及全人类产生了深远影响;且二十一世纪人口规模由最初的十六亿激增至如今七十亿以上,资源消耗和废弃物产生量增长了十倍以上,加剧了灾害发生的频率与不确定性;灾害发生造成的经济损失,生命财产损失也呈指数型增长——1995年的阪神地震、2005年的卡特里娜飓风、2008年的汶川地震,经济损失均超过1500亿美元[1],为此科学智能应对灾害的问题亟待解决。

应急预案作为应急管理处置流程的依据和基础,随着《突发事件应对法》的颁布,应急预案的编制已经成为各级政府和企事业单位应急工作的重心,然而国内应急管理领域起步较晚、基础薄弱、领域专家有限,许多中小企业在预案编制方面心有余而力不足,不是缺少相关专项,就是内容漏洞百出,若出现事故后果不堪设想,即使编制,需要的长期人力物力时间资源消费让一些企业望而却步,于是产生缺乏相关科学理论指导的畸形预案——基于编制者的个人经验或照抄照搬,产生的责任不清、措施缺乏、针对性不强等问题,使人们逐渐意识到文本式应急预案存在的种种缺陷。

二、应急预案数字化

应急预案是为针对可能发生的突发事件的事前、事发、事中、事后进行即刻管理迅速展开救援行动并降低事故损失而预先制定的具有全面性、系统性、权威性、实用性的行动方案。

国外一些发达国家早已开展应急管理信息化与应急预案数字化方面研究,并建立了较完善的应急管理信息系统,国外典型的应急预案数字化项目有美国联邦应急管理局的“e-FEMA”战略[2]系统,美国俄亥俄油气田应急响应平台[3],英国应急管理系统 IEM[4]、德国危机预防信息平台 deNIS [5]等。

国内相关领域起步较晚,基础薄弱,资源缺乏,但经过长期政

策推动支持,目前也产生了一些有代表性的应急管理项目,如国家应急平台体系[6],事故应急调度指挥过程有色非自主的 Petri 网模型的城市突发事故应急预案暨决策支持系统[7],基于 GIS 技术,结合数据、模型、方案三库,实现了城市重大事故应急决策支持,基于三维仿真技术搭建的 2008 年北京奥运会数字化消防救援预案系统[8]。

专家系统是一个包含一定领域大量专家知识与经验、可以利用人类专家的知识和方法模拟人类专家来解决问题的智能计算机程序系统,早在 20 世纪 70 年代中期,专家系统的发展已经取得了一定的成功,成为人工智能应用研究的主要领域之一。多年以后,专家系统在全世界范围内的迅速发展和广泛的使用,逐渐成为人类智能管理与决策的重要工具,本文正是讨论关于建立四方面知识库+案例推理+智能生成+决策支持的专家系统的理论可能性。

三、人工智能与应急管理

3.1 人工智能发展概况

1955 年达特茅斯会议作为人工智能普遍公认的诞生标志性事件[9],技术理论的不延伸与与各行各业的广泛结合,方便快捷高效的提高生产力,推动经济全球化,尤其是在后疫情时代,Statista 的数据库预测,2025 年全球人工智能软件市场的收入将达到 1260 亿美元,推动第三次科技革命,其展现的是系统或程序,从经验中学习和学习的能力,它与日俱增的功能和普及程度是大家都有目共睹的,

3.2 机器学习

机器学习是一种“学习算法”,无需显式编程,利用计算机程序进行经验数据的自我完善,笔者认为机器学习的核心是循环,就像编程实现的 for,中国有句古话“没人比你更聪明,只有人比你更努力”,怎么就是体现努力?那就是重复与循环,天才往往就是重复次数最多的人,机器不断地学习与重复,利用输入的经验数据进行学习,生成“学习算法”,算法性能一般情况下也与输入数据量呈负反馈循环,也就成了专家与天才,机器学习主要解决四类问题:聚类、

分类、预测和降维[10]，其主要类别有：有监督学习、无监督学习和深度学习等，在应急管理中常使用机器学习对数据进行预处理方便训练。

3.3 数据挖掘

数据挖掘涉及数据库技术、统计学和机器学习技术综合，主要分为描述性和预测性数据挖掘[11]应急管理的事前、事发、事中和事后常常伴随着各种舆论导向，有正面舆论也有负面舆论，这些负面舆论导向可能引发社会恐慌的同时造成对应急救援、应急响应的不利影响，因此舆论控制往往也对社会总体稳定和事后恢复重建有重大战略作用，数据挖掘技术就用于对负面舆论进行处理，对舆论发起人进行法律责任追究，其在事态扩大时作为处理手段运用到社交网络平台中，利用关键词技术对负面舆论进行删除和处理，对发出舆论 IP 进行捕捉，对事故责任人进行后期法律责任制裁，避免负面信息滋生，控制事态影响。

3.4 语音识别

语音识别即通过输入语音信号使得计算机识别人类语音内容含义，将语音转化为计算机可识别编码和字符文本，主流技术是基于 CNN 卷积神经网络与马尔科夫模型的声纹识别技术[11]，在事故救援和成因分析中有广泛应用前景。

3.5 计算机视觉

计算机视觉即利用算法识别由最初的一张制定图片分析潜在物品等信息到如今的视频识别技术，通过视频资料处理，分析潜在风险源与响应条件，但该技术目前一些效果仍未实现[12][13]。

四、知识库

知识库这一概念来自于两个领域的结合——人工智能的知识工程领域和传统的数据库领域，经过计算机技术的有机结合，进而产生和发展，这一概念在国内还算新颖，在企业管理中使用最多，用于降低员工培训成本，方便技术更新等，知识库的应用便捷了数据和信息的存储、重用、更新和提取，使其结构化可视[14]。

知识库系统最重要的三个模块：知识获取、知识表示和知识存储，知识表示中常用方法有机器学习中的聚类和决策树，知识表示是案例推理的基础，也是人工智能领域的重要分支，其目的是为了将各种信息数据结构化[15][16]，常用方法如人工智能应用最多的产生式表示法——许多成功的专家系统的典范，面向对象法、XML 法和机器学习中的神经网络[17]，在不同情况下用不同方法最大化表示原有数据，知识存储即用结构化形式存储非结构化数据，便于后来检索使用[18][19]。

五、案例推理 (CBR)

案例推理技术在人工智能各项领域中占据不可替代的作用，相比于规则推理的方法，案例推理机制则比较灵活，案例推理中数据获取也更为简便，推理思维更接近人脑，其基本逻辑如下：首先将获取到的新案例与案例库中案例进行对比，随机抽取得到一些属性计算匹配度，第二步，从已有案例库中选出匹配度高案例进行指导，适当时形成解决方案[20]，若匹配度未达到指定阈值，则对案例进行修改，审核后，得到处置方案，其机制一般包括以下五个过程：案例表示、匹配检索、重用、修正和保存。

5.1 5R 模型

(1) 案例表示 (Represent)：将当前案例以某种案例表征方式替换为新案例。

(2) 检索匹配 (Retrieve)：根据案例匹配算法在案例库中寻找相关案例。

(3) 案例重用 (Reuse)：选择匹配度高的案例进行重用。

(4) 案例修正 (Revise)：如果搜索出来案例与当前问题无法匹配则进行修改。

5.2 案例推理关键技术

5.2.1 案例表示

在案例推理系统中，案例表示作为数字化系统的基本内容，良

好的案例表示不仅可以数据信息以结构化形式进行存储以便于未来的重用，而且对于数据库的不断迭代更新与未来机器的学习量减小具有极大意义[22]，其表示关键在于特征选取，此领域中目前还无完全统一与最优方法，常见表示方法如：框架表示法，面向对象表示法，XML 表示法、产生式表示法，谓词逻辑表示法，语义网络表示法和基于本体的表示法等[23]。

5.2.2 案例检索

案例检索包括检索方法选取、特征权重赋值和案例之间相似度计算等方法和技术，案例检索是为了准确找到案例数据库中相似度高的案例进行匹配计算，其核心是进行案例相似度距离计算，其常见方法包括：最近邻策略、决策树检索、知识引导策略和模板检索策略[24]。

5.2.3 案例重用

通过案例检索，从案例数据库中得到了与目标案例相似度较大的案例集合，如果目标案例与案例数据库中历史案例相似度达到一定阈值，则可考虑重用，若未达到设定阈值，则会进行案例修正 [25]。最大相似度重用即选取相似度最大的历史案例进行处置方法的重用，多数重用即选取相似度较大的一系列案例进行选择，结论占多数的则作为解决方案。

5.2.4 案例修正

通过案例检索可以得到的相似度较大的案例集合，通过两种方法最大相似度和多数重用原则选择时，如果此历史案例相似度与所需解决问题案例相似度达到阈值，则我们可以直接使用，如果没有达到此设定阈值，则需要案例修改，案例修改后存在仍无法近似匹配的问题，但不会删除，而是保存到案例库中作文新的数据信息，案例修改涉及特定领域，其技术也是案例推理的难点之一[26]。

转换法：基于案例进行模块组合和修改调整，改变案例数据信息值和结构。

派生重演法：基于以前案例的修改属性进行学习，为以后的修改意见提供指导。

组合调整：将相似度较大案例集合有机组合解决部分。

5.2.5 案例保存

案例库系统具有机器学习的效用，在于案例保存机制——将检索出相似度达到阈值的案例进行保存，用来扩充原有案例库，因此案例库的不断增长，可以为新案例问题的解决提供较全面的处置方案，而且也不会带来检索困难的问题，其在于案例保存机制会自动选择具有代表性的案例进行检索候选，去除冗余数据，提高检索精度[27]。

六、案例生成算法研究

6.1 案例知识表示

6.1.1 知识获取与数据处理

数据获取方面，我们首先要保证专业性，以至于不会出现常识性错误，我们会从专业网站网址，以煤矿事故为例，其相关官网如中国煤矿生产网等等，然后我们要保证数量充足，以至于不会出现欠拟合，我们使用爬虫软件爬取大量案例信息，基于 python 的 urllib 库可提供一系列 URL 功能，BeautifulSoup 用于解析网页，Python-docx 用来 Word 文档生成[28]。

(1) 数据量化表示

使用案例推理技术，将所待生成案例处置方案与历史案例库对比相似，根据特征提取计算不同案例间相似度，为方便计算常常需要将数据属性值量化。

(2) 缺失值处理

搜集数据往往存在各种因素使得数据存在缺失值，这常常导致数据缺失或使得结果计算不准确与模型性能不优，一般处理办法如：填充法，删除法和模型预测法。

(3) 失衡数据处理

失衡数据即案例数据集中某些特征值的数量极度少于其他案例

数据集中该特征值的数量,而最终导致生成案例指导性欠佳。常用解决方法如:使用平均或内插补充法,对较多特征值样本进行删减等。

6.1.2 案例知识表示

人工智能即机器模拟人脑运作方式,进行问题的解决和处理,上世纪七十年代美国学者 Minsky 提出人类思维框架理论,即人类在遇到问题时的解决方案包括回想曾经遇到同种问题的解决方案,但人类不能拥有强大的记忆以至于存储所有问题解决方案并迅速找到最优解,但是机器可以做到,并且不断更新与迭代方案库。

框架法常常需要对每一个案例都有唯一性标识,因而需要对每个案例提供标号和名称。案例框架有许多“槽”,每个“槽”有许多“侧面”,每个“侧面”有自己对应侧面值。在框架法的案例表示中,我们以煤矿事故案例框架法表示为例:

事故案例编号	0001		
框架名称	煤矿事故案例框架		
槽 1	煤矿企业概况		
槽 2	事故基本特征		
槽 3	事故描述		
槽 4	事故原因分析		
槽 5	事故应急资源		
槽 6	应急机构及职责		
槽 7	事故处置效果		
侧面 2.5	瓦斯事故类型	瓦斯燃烧	枚举型
		瓦斯突出	枚举型
		煤与瓦斯突出	枚举型
		地表水水害	枚举型
		矿层顶板水害	枚举型
侧面 2.6	水灾事故类型	矿层底板水害	枚举型
		老空积水水害	枚举型
		地质构造水害	枚举型
		工业用水水害	枚举型
		外因火灾	枚举型
侧面 2.7	火灾事故类型	内因火灾	枚举型

主要依据以上构造进行案例数据存儲。

6.2 案例检索

案例检索即从历史案例库中寻找具有代表性的案例属性,作为案例检索的特征,依据这些特征相似度计算方法得出案例相似度最大的案例进行处置方法提取以便于当前案例问题的处理,因此高效、科学的案例检索算法尤为重要[30]。

6.2.1 三级案例检索

我们采用三级检索法[31],第一级就是上文说到的根据事故类型:水灾、火灾和瓦斯爆炸,这样将减少近一半的数据量,第二步采用最近邻算法根据事故特征中的事故地点、应急响应级别和事故类型,第三步是采用最近相邻算法,根据事故特征的企业概况、事故描述中的特征选取相似度达到阈值(0.8)的案例进行检索,大大提高检索效率。

6.3 案例相似度计算^[31]

局部相似度

同样以煤矿事故为例,将案例数据属性分为三种类型:文本、数值和枚举。

(1)数值型。是指案例特征值是连续型或者离散型的数值,能够准确地用数字来表示事故案例的属性特征。例如灾害中的死伤人、经济损失等。计算这类特征相似度的方法如下。

$$sim_x(i, j) = 1 - \frac{|c_i^x - c_j^x|}{c_{max}^x - c_{min}^x}$$

$sim_x(i, j)$ 代表新案例 i 和已有案例 j 在特征 r 上的局部相

似度, c_{max}^x 是特征 r 的最大值, c_{min}^x 是特征 r 的最小值; c_i^x 是新案例 i 的第 r 个属性值, c_j^x 是历史事故案例中案例 j 第 r 个属性值。

(2)文本型。是采用字段匹配的方法,各特征值之间是互不相关的。如煤矿灾害的事故类型、矿井开采方式、通风方式以及企业性质等。该类特征的局部相似度计算方法如下。

$$sim_x(i, j) = \begin{cases} 1, C_i^x = C_j^x \\ 0, C_i^x \neq C_j^x \end{cases}$$

当新案例 i 的第 r 个特征和历史案例库中案例 j 的第 r 个特征一样时,该局部相似度的值为 1,相反该值为 0。

(3)枚举型。能够按照灾害不同指标的程度制作评估等级。这类特征的局部相似度计算方法如下。

$$sim_x(i, j) = 1 - \frac{|C_i^x - C_j^x|}{N}$$

C_i^x 是新案例 i 第 r 个属性值, C_j^x 是历史案例库当中案例 j 第 r 值; N 是枚举类型中属性的最大值。

案例总体相似度计算

案例总体相似度计算即各局部相似度与各权重的乘积的和,计算公式如下。

$$sim_x(i, j) = \sum_{x=1}^n W_x sim(i, j)$$

其中 W_x 是特征 r 计算相似度是权重。

七、决策支持模型

我们的知识库+案例推理+生成算法已经可以做到根据不同事故或企业情况生成数字预案并存入案例库中,但是真正的突发事件到来时,很可能出现预案制定内容之外甚至是历史案例库之外的情况,因此需要一种知识表达方式供计算机理解,以应急预案在决策中所利用的技术,给人类提供智能分析,这就是数字化生成之外的决策支持系统,统称为专家系统。

基于 Web 和 XML 语法结构的本体描述语言,可以很好进行本体建模处理,在基于 Web 的本体描述语言中,常用 OXL 和 SWRL 相结合应用于应急决策知识建模,斯坦福大学开发的基于 JAVA 的 Protégé [35] 本体编辑与知识获取软件,可对中文支持进行本体构建,其 4.3 版本可支持 SWRL 语义规则进行本体编辑,刘晓慧建立的 PEOMG 本体模型,可以为应急预案流程的执行过程提供智能化支持。

八、结语

应急管理领域我国起步相对较晚,但面对共同的灾害,我们如何像工业革命一般用短短几年走完发达国家几十年的道路,多快好省的达到国家总体安全观的治理能力新高度,唯有顺应当下潮流,“互联网+数字化”将应急治理由“人治”转变为“人与智”,将各种损失降到最低,全面保障人民的美好生活需要。

参考文献:

- [1]张继权 《中国灾害风险管理导论》
- [2]黄以宽.e-FEMA 对我国应急管理体系建设的启示.信息化建设.2006(6).
- [3]张超,裴玉起,邱华.国内外数字化应急预案技术发展现状与趋势[J].中国安全生产科学技术.2010,6(5):154-158.
- [4]李雪峰,2010.英国应急管理体系、方法与借鉴.北京市:国家行政学院出版
- [5]邵瑜.德国的危机预防信息系统[J].信息化建设,2005(8):46-48.
- [6]范维澄,陈涛,国家应急平台体系建设现状与发展趋势[J].中国突发事件防范与快速处置优秀成果选编,2009:3.