



工业和信息化部“十二五”规划专著

消费品质量安全信息分析 ——理论、方法与技术

XIAOFEIPIN ZHILIANG ANQUAN
XINXI FENXI

王 理 许应成 著



北京航空航天大学出版社
BEIHANG UNIVERSITY PRESS

014038726

F273.2
647



工业和信息化部“十二五”规划专著

消费品质量安全信息分析 ——理论、方法与技术

王理 许应成 著



F273.2
647

北京航空航天大学出版社



北航

C1726312

010387530

内 容 简 介

本书从国内外消费品质量安全信息分析研究现状出发,结合我国消费品质量安全实际情况,着重论述了消费品质量安全信息分析的理论、方法与技术。全书共分10章,主要包括消费品质量安全分析基础知识、消费品质量安全信息采集与集成、基于话题跟踪的消费品质量安全信息分析、基于文本特征的消费品质量安全Web文本分类、消费品安全事件中的事件信息抽取与分析,以及消费品质量安全信息监测及预警等。

本书可以作为高等学校管理科学、信息科学和质量工程等专业的教材或参考书,同时也可供消费品质量安全信息分析相关的工程技术人员、研究工作者和管理人员阅读和参考。

图书在版编目(CIP)数据

消费品质量安全信息分析:理论、方法与技术/王理,许应成著. —北京:北京航空航天大学出版社, 2014.4

ISBN 978-7-5124-1518-8

I. ①消… II. ①王… ②许… III. ①消费品—质量管理—安全管理—信息处理 IV. ①F273.2

中国版本图书馆CIP数据核字(2014)第050042号

版权所有,侵权必究。

消费品质量安全信息分析——理论、方法与技术

王 理 许应成 著

责任编辑 宋淑娟 刘晓明 王 实
张冀青 周华玲

*

北京航空航天大学出版社出版发行

北京市海淀区学院路37号(邮编100191) <http://www.buaapress.com.cn>

发行部电话:(010)82317024 传真:(010)82328026

读者信箱: goodtextbook@126.com 邮购电话:(010)82316524

北京兴华昌盛印刷有限公司印装 各地书店经销

*

开本:787×1092 1/16 印张:10.5 字数:269千字

2014年4月第1版 2014年4月第1次印刷 印数:1000册

ISBN 978-7-5124-1518-8 定价:32.00元



前 言

消费品质量安全关系国计民生,不仅备受消费者关注,而且是社会舆论和各国政府关注的焦点。我国正处在社会和经济发展的重大战略机遇期,经济飞速发展。但是,大部分企业以产品符合国家标准和法规要求为目标,追求消费品质量安全的符合性;而对于我国标准中未规定的安全要求,没有采用国际上已经普遍认同的“合理可预见的误用情况下产品应该是安全的”理念对产品的安全性进行深入研究,因而表现出产品虽然符合标准要求,但是仍然存在安全隐患的问题。消费品质量安全信息是消费品质量安全监控的源头,只有动态、实时和有效地获取消费品质量安全信息,深入分析质量安全信息反映的消费品质量安全问题,才能准确把握消费品质量安全状况,及时发现系统性、区域性和行业性消费品质量安全问题。从消费品质量安全风险监测与信息分析技术研究来看,我国尚未建立系统的质量安全信息监测和信息分析的技术支撑体系,缺乏对消费品质量安全信息传播规律的掌握,缺少信息获取的技术手段和信息分析的方法等。这些导致了我国消费品质量管理缺乏必要的科学手段,以及产品安全风险监控技术应用范围狭窄和共性技术缺失等问题。因此我国亟须系统地开展消费品质量安全信息分析及关键技术的研究。

本书以循序渐进的方式介绍消费品质量安全信息分析的相关理论、方法与技术。在内容安排上侧重于以消费品质量安全信息采集、推理、分析与处置的全生命周期为主线展开论述,并且尽量在每一章节都添加一些应用实例,用以辅助说明,力求使读者在阅读本书后即可掌握与消费品质量安全信息分析相关的理论、方法与技术。在叙述方式上力求由浅入深,在着重介绍技术与方法的同时,也说明能够有效使用这种方法的条件和理由。

本书的内容分为10章。第1章全面阐述消费品质量安全问题。第2章综述消费品质量安全信息分析过程中使用的理论、方法和技术。第3~9章主要对消费品质量安全信息生命周期模型中涉及的关键技术进行探讨:第3章论述消费品质量安全信息采集与处理,它是后面几章的基础;第4章探讨基于话题跟踪的消费品质量安全信息分析;第5章研究基于文本特征的消费品质量安全Web文本分类;第6章重点探讨消费品质量安全事件中的时间信息抽取与分析;第7章论述消费品质量安全事件中的地理信息抽取与分析;第8章研究基于贝叶斯网络的消费品质量安全信息推理;第9章论述基于支持向量机的消费品质量安全预警模型。第10章对消费品质量安全信息的应用进行论述,其中主要研究消费品质量安全信息监测与预警平台。

许应成负责第1、3、8~10章的撰写,王理负责第2、4~7章的撰写,最后由王



理统稿。

本书可以作为高等学校管理科学、信息科学和质量工程等专业的教材或参考书,同时也可供消费品质量安全信息分析相关的工程技术人员、研究工作者和管理人员阅读和参考。

本书主要内容来源于作者主持的国家自然科学基金和国家科技支撑计划等课题的研究成果,另外一些内容是参阅国内外有关文献撰写而成的。作者在此对直接或间接对本书写作有帮助或贡献的书籍或文献的作者表示诚挚的感谢。

此外,本书的写作和出版得到了国家自然科学基金项目(71301152, 71271013, 70971005, 91224007)、“十二五”国家科技支撑计划课题(2013BAK04B02)、中国博士后科学基金项目(2012M520008, 2013T60091)、国家社会科学基金项目(11AZD096)、国家留学基金项目(2011307460)、北京市自然科学基金项目(9142012)和北京市科技项目(Z121100000312018)等课题的支持。邓卫国、祁静、王思远、钱中、向体燕、蔡华利、施慧斌、潘守慧、杨跃翔、孟德运、王君美、于珊、谢俊、郝乙和刘志莲等对本书的写作做出了大量贡献,北京航空航天大学出版社为本书的顺利出版做了大量细致的工作,作者对此表示衷心的感谢。

由于作者水平有限,书中难免存在不足之处,敬请广大读者和同行专家不吝指正。

作者

2014年3月



目 录

第1章 绪 论	1
1.1 消费品质量安全信息分析的重要性	1
1.1.1 研究背景	1
1.1.2 研究意义	4
1.2 消费品质量安全信息分析的研究现状	5
1.2.1 国外研究现状	5
1.2.2 国内研究现状	9
1.3 章节安排	10
参考文献	11
第2章 Web 挖掘理论	12
2.1 概 述	12
2.1.1 数据挖掘	12
2.1.2 Web 挖掘	14
2.2 Web 挖掘的分类	15
2.2.1 Web 内容挖掘	15
2.2.2 Web 结构挖掘	17
2.2.3 Web 使用挖掘	17
2.3 Web 挖掘算法	19
2.4 Web 信息抽取技术	21
2.4.1 MUC 和 ACE	21
2.4.2 TAC 和 KBP	22
2.4.3 Web 信息抽取关键技术	24
2.5 基于 Web 挖掘的网络舆情智能分析	26
2.5.1 网络舆情信息分析处理	26
2.5.2 基于社交网络的信息传播研究	26
2.6 消费品质量安全信息分析和预警	27
参考文献	28
第3章 消费品质量安全信息采集与集成	35
3.1 消费品质量安全信息采集及处理	35
3.1.1 概 述	35
3.1.2 消费品质量安全信息采集	36
3.1.3 消费品质量安全信息处理	39



3.2 消费品质量安全信息要素提取·····	40
3.2.1 消费品质量安全影响因素·····	40
3.2.2 信息结构·····	42
3.3 消费品质量安全信息集成·····	46
3.3.1 集成模型·····	46
3.3.2 基于本体的数据集成过程·····	48
参考文献·····	49
第4章 基于话题跟踪的消费品质量安全信息分析 ·····	50
4.1 话题跟踪研究现状·····	50
4.1.1 相关概念·····	50
4.1.2 研究现状·····	51
4.2 质量安全事件 Web 新闻话题跟踪模型构建·····	52
4.2.1 Web 文档预处理·····	52
4.2.2 特征选择和文本表示·····	52
4.2.3 特征权重计算·····	53
4.2.4 Single-Pass 增量聚类算法·····	54
4.3 改进的相似度计算方法·····	55
4.3.1 质量安全事件发生时间抽取·····	55
4.3.2 考虑时间距离的相似度计算方法·····	57
4.4 实验分析·····	57
4.4.1 评价指标·····	57
4.4.2 结果分析·····	58
参考文献·····	60
第5章 基于文本特征的消费品质量安全 Web 文本分类 ·····	63
5.1 文本分类研究现状·····	63
5.2 Web 文本分类关键技术研究·····	65
5.2.1 文本预处理·····	65
5.2.2 文本特征表示·····	66
5.2.3 特征降维·····	67
5.3 改进的基于 VSM 的特征项权重计算方法·····	69
5.3.1 传统的特征项权重计算方法·····	69
5.3.2 改进的特征项权重计算方法·····	69
5.4 基于规则和 KNN 的三层架构的分类器构造·····	70
5.5 实验分析·····	72
5.5.1 评测指标·····	72
5.5.2 实验数据及过程·····	72
5.5.3 结果分析·····	74



参考文献	76
第 6 章 基于网络案例的消费品质量安全时间信息分析	78
6.1 时间信息基本概念及研究现状	78
6.1.1 时间信息基本概念	78
6.1.2 时间信息抽取研究现状	78
6.1.3 时间信息抽取的意义	79
6.2 时间关系理论	79
6.3 网络案例时间表达特征	80
6.3.1 网络案例时间构成	80
6.3.2 网络案例时间位置特征及常用词分析	82
6.4 网络案例发生时间提取方法	82
6.4.1 时间用词到时间戳的映射	82
6.4.2 基于向下覆盖的发生时间提取方法	83
6.4.3 实验分析	84
6.5 案例时序分析	85
参考文献	87
第 7 章 基于规则推理的消费品质量安全地理信息分析	89
7.1 地名实体抽取相关研究	89
7.1.1 命名实体识别方法	89
7.1.2 中文地名的特点	89
7.1.3 消费品质量安全事件地名实体抽取的意义	90
7.2 地名实体定义及表达特征分析	90
7.2.1 地名实体的界定	90
7.2.2 地点实体信息的表达特征	90
7.2.3 现有分词系统处理地名实体的不足	92
7.3 网络案例地名实体的识别方法	92
7.3.1 基于规则推理的地名实体识别流程	92
7.3.2 网络案例发生地点的分析和判断	94
7.3.3 实验分析	95
7.4 案例地点分析	96
参考文献	97
第 8 章 基于贝叶斯网络的消费品质量安全信息推理	98
8.1 贝叶斯网络推理与故障树分析法研究现状	98
8.1.1 贝叶斯网络推理研究现状	98
8.1.2 故障树分析法研究现状	99
8.2 案例知识发现的过程及特征分析	100



8.2.1	消费品质量安全案例的特征分析	100
8.2.2	案例知识发现的过程	101
8.3	多实体贝叶斯网络建模	101
8.3.1	故障树常规逻辑关系向多实体贝叶斯网络的转化	102
8.3.2	开关事件向多实体贝叶斯网络的转化	102
8.3.3	共因事件向多实体贝叶斯网络的转化	104
8.3.4	人因事件向多实体贝叶斯网络的转化	105
8.4	多实体贝叶斯网络参数学习算法	107
8.4.1	现有的参数学习算法	107
8.4.2	在线最大期望学习算法	107
8.4.3	基于证据理论的多实体贝叶斯网络参数学习算法	108
8.5	多实体贝叶斯网络系统安全评价	109
8.5.1	后果概率	109
8.5.2	重要度	110
8.6	实验分析	110
	参考文献	114
第9章	基于支持向量机的消费品质量安全预警模型	116
9.1	支持向量机算法	116
9.1.1	支持向量机算法概述	116
9.1.2	支持向量机的数学模型	117
9.1.3	支持向量机算法的应用	118
9.2	特征加权支持向量机模型	120
9.2.1	特征加权支持向量机的算法	120
9.2.2	特征加权支持向量机的步骤	122
9.3	消费品质量安全风险预警模型的构建	123
9.3.1	玩具类消费品安全影响因素分析	123
9.3.2	玩具类消费品安全风险预警指标选择	125
9.3.3	玩具类消费品质量安全预警流程	128
9.3.4	消费品质量安全风险预警模型的建立	130
9.3.5	核函数选择	131
9.3.6	模型的评价	132
9.4	实验分析	132
9.4.1	数据收集及预处理	132
9.4.2	预警应用实验	136
9.4.3	预警结果分析	139
	参考文献	140



第 10 章 消费品质量安全信息监测及预警平台	143
10.1 平台建设的目标与原则	143
10.1.1 平台建设的目标	143
10.1.2 平台建设的原则	143
10.2 平台的基本结构	144
10.3 信息采集	145
10.3.1 论坛、博客、网页的采集	146
10.3.2 搜索引擎采集	148
10.3.3 案例信息采集	149
10.4 决策支持系统	150
10.4.1 信息查询模块	150
10.4.2 统计分析模块	152
10.4.3 数据挖掘模块	155
参考文献	156
结 论	157

消费品是指居民日常生活中所必需的产品,不包括国家专门部门或机构所管理的药品、药品、化妆品、特种产品、危险品、易燃易爆品、军用产品等。目前,消费品可分为 12 类,主要包括建筑材料、服装鞋帽及配件、儿童用品、家具及配件、装饰装修材料、文教体育办公用品、美容护肤护发产品、日用包装物及容器、厨房厨房用消费品、通信设备及相关产品、信息系统安全产品、电子计算机及相关产品、家用电器、工艺品及其他制造产品、灯具等照明光源、交通运输工具、日用杂品等。

“消费品质量安全”是指消费品在正常使用或通常预期的用途方面符合“一般安全规定”(general safety requirement)。“一般安全规定”描述了确保消费品合乎合理的安全程度的要求。只要有任何我国认可的标准适用于某消费品,而该消费品符合该标准,则该消费品即视为符合一般安全规定。

随着我国经济和社会的发展和进步,我国对消费者权益的保护程度也不断提高,对消费品的“合理的安全程度”提出了更高的要求,已经由过去的只关注保护消费者人身健康和财产安全向保护公众安全和生态安全扩展。确定消费品是否合乎“合理的安全程度”,应综合考虑对消费者保护的力度,以及改善消费品安全程度的成本、可能性,是否有合理的方法来使得消费品更为安全等问题。随着生活水平的不断提高,消费者对消费品不再局限于基本功能的满足,而更加强调消费过程的安全性;然而,社会的快速发展带来了更多的未知隐患,使得消费者在消费过程中面临更多的不确定性。

1. 我国消费品质量安全形势严峻

(1) 最近几年,一系列消费品质量安全事件频繁发生,消费品质量安全问题日益突出。例如,2007 年中国玩具小型可拆卸部件造成儿童窒息导致的“美泰召回”事件,2008 年“汽车召回品中有马酸二甲酯导致严重过敏”事件,2009 年“婴儿纸尿裤含有石棉致癌物”事件,2010 年“饰品中镍超标”事件,2011 年“婴儿奶瓶中的双酚 A”事件、“玩具中苯二甲酸酯”事件,2012 年“大众汽车 DSG 变速箱召回”事件,2013 年“毒校服”事件、“苹果 515 后盖门”事件。

第1章 绪论

在人类社会生活中,消费品已成为必不可少的物质资源,消费品质量安全问题也日益成为人们关注的热点。随着互联网技术的发展,特别是微信、博客、维基、微博、论坛、社交网络和内容社区等新兴社会化媒体的出现,消费品质量安全信息在采集、发布、评估、预警和处置等方面越来越复杂化和专业化;同时消费品质量安全信息在信息来源、信息类型、网络结构、文本特征、传播渠道等方面也出现了很大变化。因此,科学、合理地消费品质量安全信息进行分析,并对相关的理论、技术和方法进行研究,具有十分重要的理论价值和实际意义。

1.1 消费品质量安全信息分析的重要性

1.1.1 研究背景

根据 GB/T 22760—2008《消费品安全风险评估通则》的定义,消费品指为满足社会成员生活需要而销售的产品。不包括由专门部门进行特殊监管的食品、药品、化妆品、特种设备、航空器具、船舶、军用产品等。目前,消费品可分为 17 类,主要包括纺织制品、服装鞋帽及附件、儿童用品、家具及配件、装饰装修材料、文教体育办公用品、美容护肤护发产品、日用包装物及容器、卫浴厨房用消费品、通信设备及相关产品、信息系统安全产品、电子计算机及相关产品、家用电器、工艺品及其他制造产品、灯具等照明光源、交通运输工具、日用杂品等。

“消费品质量安全”是指消费品在指定用途或通常预期的用途方面符合“一般安全规定”(general safety requirement),“一般安全规定”描述了确保消费品合乎合理的安全程度的要求。只要有任何我国认可的标准适用于某消费品,而该消费品符合该标准,则该消费品即视为符合一般安全规定。

随着我国经济和社会的发展和进步,我国对消费者权益的保护程度也不断提高,对消费品的“合理的安全程度”提出了更高的要求,已经由过去的只关注保护消费者人身健康和财产安全向保护公众安全和生态安全扩展。确定消费品是否合乎“合理的安全程度”,应综合考虑对消费者保护的程度,以及改善消费品安全程度的成本、可能性、是否有合理的方法使得该消费品更为安全等问题。随着生活水平的不断提高,消费者对消费品不再局限于基本功能的满足,而更加强调消费过程的安全性;然而,社会的快速发展带来了更多的未知隐患,使得消费者在消费过程中面临更多的不确定性。

1. 我国消费品质量安全形势严峻

(1) 最近几年,一系列消费品质量安全事件频繁发生,消费品质量安全问题日益突出

例如,2007 年中国玩具小型可拆卸部件造成儿童窒息导致的“美泰召回”事件,2008 年“皮革制品中富马酸二甲酯导致严重过敏”事件,2009 年“婴儿爽身粉含有石棉致癌物”事件,2010 年“饰品中镉超标”事件,2011 年“婴儿奶瓶中的双酚 A”事件、“玩具中邻苯二甲酸酯”事件,2012 年“大众汽车 DSG 变速箱召回”事件,2013 年“毒校服”事件、“苹果 315 后盖门”事件、



铜水龙头含铅超标可导致铅中毒事件,等等。这些消费品质量安全事件涉及行业广泛、产品量大和面广,且屡禁不止,造成了严重的人身伤害和经济损失,引发了广大人民群众对政府消费品质量安全监管的质疑,积累了社会的不满情绪,不利于保护民生与和谐社会的构建;同时,这些消费品质量安全事件引起了一系列关于“中国制造”的讨论,中国品牌被部分媒体随意解读,对相关行业持续健康发展产生了严重的负面影响。

来自国家质量监督检验检疫总局的数据显示,在国内制造业产品方面,2013年国家监督抽查共抽查了17 020批次产品,批次抽样合格率为88.9%,同比下降0.9个百分点,大、中、小型企业批次抽样合格率分别为95.3%、92.0%和86.3%,与2012年相比均有不同程度下降。存在的主要问题包括:电器设备方面,部分自动电饭锅、电磁灶、厨房机械和商用电热食品加工设备的电气安全指标不合格;汽车配件方面,部分GPS导航产品、灯具的配光性能、传动带的尺寸、制动软管和制动器衬片不符合标准要求;儿童用品方面,玩具的小零件和尖端项目不合格问题依然存在,部分童装纤维成分、pH值等项目不合格,部分童鞋的游离甲醛、皮革中六价铬、耐磨性能等项目不符合标准的规定;建材方面,部分刨花板、浸渍纸层压木质地板、竹地板、胶合板和细木工板的甲醛释放量超标,部分地坪涂装材料的有害物质限量不合格,部分门、窗用未增塑聚氯乙烯型材的拉伸冲击强度不达标,部分建筑防水卷材的热老化项目不合格,部分新型墙体材料的节能性能指标不合格。在产品质量投诉方面,2013年全国质检系统12365机构共受理业务咨询、举报和投诉业务总量达82.5万起,其中举报6万起,投诉3.98万起。投诉举报排在前5位的产品分别是非酒精饮料、肉制品、手机、水泥及其制品、豆制品。在全国质检系统受理的消费者对产品质量的申诉中,申诉量排在前5位的产品依次为食品、家具、轿车、移动电话、电视。以轿车和移动电话为例,轿车质量问题集中出现在汽车发动机、变速箱、刹车系统、减震器等主要部件上,移动电话的主要问题有电池容量不达标、屏幕显示异常、触摸屏不灵敏、网络信号时有时无、充电器漏电等。质检数据显示,我国消费品质量安全问题较为严重,缺少必要的安全信息监测机制,消费品质量安全隐患亟待解决。

(2) 我国出口消费品由于质量安全问题被大量召回或退回,经济损失巨大,严重影响相关产业发展

根据美国消费品安全委员会(Consumer Product Safety Commission, CPSC)发布的2010年度产品安全报告,中国大陆产品位居美国通报召回产品首位,占通报产品总量的60.6%;在2011年通过欧盟RAPEX(the Rapid Alert System for Non-food Consumer Products)系统发布的所有通报中,有54%(即839项)通报产品的原产国为中国(含香港);在2012年通过RAPEX系统发布的所有通报中,有58%(即1 126项)通报产品的原产国为中国(包括香港)。根据有关统计,2004—2010年,我国几乎每年都是世界上受通报最多的国家,每年由此造成的直接和潜在的经济损失约500亿美元。据调查测算,我国出口企业中近60%遭到国外技术性贸易措施制约,对我国出口产品造成的影响每年超过450亿美元,占年出口总额的25%以上。从个体消费品质量安全事件看,每次产品召回都导致了严重的经济和社会损失。以“美泰召回”事件为例,导致全球召回中国产玩具1 820万件,给每年出口创汇达20多亿美元的中国玩具业造成严重影响,导致1 800多家玩具企业倒闭(王赞松等,2007,2009)。

(3) 消费品质量安全问题造成严重个体或群体伤害,极易引发系统性的社会风险
玩具小部件脱落导致的儿童窒息、婴幼儿纺织品有害化学物质超标导致的皮肤过敏、电器附件绝缘失效导致的火灾等屡见报道,造成严重的人身伤害和财产损失。据报道,我国每年因



电器安全问题导致的人身伤害和死亡在3 000人以上,每年因装饰材料和家具安全问题引发的中毒人数在10万人以上。据2009年欧盟伤害报告,欧盟每年因发生与产品安全有关的事故死亡人数在5 700人以上;“家庭和休闲”事故造成的非致命伤害占欧盟医院治疗伤害患者的74%,其中50%的“家庭和休闲”事故伤害与产品有关。根据澳大利亚有关产品伤害数据统计研究,消费品质量安全问题导致的死亡、需住院治疗和需急诊治疗的人数比例约为1:40:1 000,还有大量的轻微伤害未到医院治疗处理,监测到的消费品质量安全伤害只是冰山一角。由此可见,消费品质量安全问题如不能及时发现和处理,有可能带来严重的群体性伤害问题,极易引发系统性的社会风险。

2. 党中央和政府高度重视消费品质量安全

消费品质量安全问题不仅影响消费者人身健康和经济的健康发展,也直接影响公众对政府部门的信心,直接危及公共安全和社会稳定,受到社会各界的普遍重视和关注。

2011年3月,党中央和政府部门在第十一届全国人民代表大会第四次会议上通过的《国民经济和社会发展第十二个五年规划纲要》中明确提出要“创新社会管理体制机制,加强社会管理能力建设,建立健全中国特色社会主义社会管理体系,确保社会既充满活力又和谐稳定。”明确要求“推动建立主动防控与应急处置相结合、传统方法与现代手段相结合的公共安全体系”。从内涵上看,消费品质量安全问题关系到经济发展和社会和谐稳定,开展消费品质量安全风险监控工作是创新社会管理机制,落实《国民经济和社会发展第十二个五年规划纲要》的重要内容。

《国家中长期科学和技术发展规划纲要(2006—2020年)》将公共安全列为重点发展领域,并提出加强对突发公共事件快速反应和应急处理的技术支持、提高早期发现与防范能力、增强应急救护综合能力以及加快公共安全装备现代化等要求。消费品的使用与消费安全,涉及消费者人身安全和财产利益,涉及国家和地区经济的健康发展,涉及社会的和谐与稳定,消费品安全管理已成为国家公共安全管理的重要组成部分。关注消费品质量安全就是进一步加强公共安全发展,提高公共安全事件快速反应的应急处理能力。

党的十八大报告中提出:“必须以保障和改善民生为重点,解决好人民最关心、最直接、最现实的利益问题,努力让人民过上更好生活。”消费品质量安全直接关系到消费者生命以及财产等安全,切实从理论、方法和技术方面解决消费品质量安全问题就是解决好人民最关心、最直接、最现实的利益问题。

国务院文件《质量发展纲要(2011—2020年)》中就重点消费品提出增强消费品质量安全溯源能力,建立质量安全联系点制度,健全质量安全监管长效机制。国家质检总局文件《质检事业发展“十二五”规划》要求以从严治检、强化监管为着力点,不断完善中国特色质检工作体系,加强安全信息监测,对消费品质量安全事件要做到有源可溯,将安全事件控制在源头上,并提出“抓质量、保安全、促发展、强质检”的质检工作方针,要求进一步做好安全信息监测工作,努力做好有害物质测试方法研究,完善测试方法标准体系。

目前,发达国家的消费品质量安全监控体系、理论与方法比较完善。我国虽然从国家层面上为消费品质量安全提升制定了规划和发展目标,但在消费品质量安全监测理论、体系和技术方面的研究相对缺乏。近年,我国消费品质量安全风险监控刚开始起步,涉及领域窄,还未进行系统化研究。从消费品质量安全风险监控与信息分析技术研究来看,我国尚未建立系统的



质量安全风险信息监测和信息分析的技术支撑体系,因此,我国亟须系统地开展消费品质量安全信息分析及关键性技术研究。

1.1.2 研究意义

当前,消费品质量安全问题频发,不仅影响到消费者的人身安全和经济的健康发展,而且也直接影响到公众对政府部门的信心,直接危及公共安全和社会稳定,受到了社会各界的普遍重视和关注。因此,对消费品质量安全信息进行深入分析,及时发现有可能造成系统性社会问题的重大风险,遏制消费品质量安全事件频发的状况已经刻不容缓。针对消费品质量安全的特性,主动实施风险监控预警来降低系统性风险,实现由“后端监管”向“前端预防”的转变,是我国现阶段消费品质量安全管理的一条重要途径,也是创新社会管理的一个重要组成部分。

信息资源是消费品质量安全监控和预警分析的重要依据。在当今网络化信息时代,除了传统的制度性渠道之外,非制度性渠道中的 Web 信息资源汇集了各类事件的新闻报道和舆论评价,是消费品质量安全事件的重要信息平台,对事件信息管理与控制具有广泛作用。互联网信息常被作为消费品质量安全事件研究的信息载体。一方面,互联网为消费品质量安全事件信息的产生、传播和扩散提供了平台,使得信息的发布、阅读和传播极为便利,形成了主流新闻媒体和个人新闻网站并存的多样化信息获取来源;另一方面,由于与特定主题相关的质量安全信息往往零散、孤立地分布在不同的时间段和信息源,并且数据往往是异构的,这给消费品质量安全事件信息的挖掘分析带来了极大困难。此外,消费品质量安全信息具有种类繁多、分布范围广、数据量大、变化快等特点,而且消费品质量安全事件信息在早期阶段极其隐蔽,一旦到达某个时点,则传播速度极快,转折点难以捕获,由此带来的问题难以用传统的统计分析方法解决。因此,迫切需要对消费品质量安全 Web 信息分析和挖掘技术进行研究,构建信息监测和预警系统。然而,目前我国政府部门在消费品质量安全事件的 Web 信息融合与分析 and 风险监控方面,仍处于以手工方式为主,以智能计算为辅的低水平状态,先进的 Web 信息分析技术在我国消费品质量安全管理体系中尚未得到充分应用。迄今为止,我国还没有一个高度集成、智能、自动化的消费品质量安全信息公共服务平台对这些信息进行采集、加工、集成和深度挖掘,导致相关信息不能及时共享,进而影响到消费品质量安全的实时监测、适时预警和快速响应,严重制约了我国消费品质量安全的管理水平。

本书以 Web 上的消费品质量安全信息为研究对象,利用 Web 挖掘等理论方法研究 Web 上的质量安全信息,讨论了质量安全 Web 文本多层次分类方法,构建了质量安全事件 Web 新闻话题跟踪模型,讨论了质量安全个性化信息服务用户兴趣建模方法,建立了基于关联规则和基于贝叶斯推理的消费品质量安全因子挖掘分析方法。本书相关章节的研究内容可为消费品质量安全信息的智能分析处理和深度挖掘提供理论基础,也为今后建设一个快速、高效、智能的消费品质量安全信息公共服务平台提供技术支撑,同时也可为研究消费品质量安全事件的传播规律提供帮助,这将有助于推动我国消费品质量安全信息管理工作向数据驱动模式方向发展,从而有助于提高消费品质量安全监控和预警决策的科学性。

本书研究的多种 Web 挖掘理论方法也是近几年信息技术领域的研究热点和难点,本书提出的质量安全 Web 文本分类方法、话题跟踪模型、个性化信息用户兴趣建模等理论方法不仅可用于消费品质量安全领域 Web 信息的挖掘和分析,而且在修改和扩展之后,还可用于网络舆论监控等其他领域。



1.2 消费品质量安全信息分析的研究现状

1.2.1 国外研究现状

目前,国外一些发达国家在消费品质量安全信息监测与分析方面开展了大量富有成效的工作,也取得了显著效果。其中美国、欧盟、日本等起步较早,积累了很多可供借鉴的经验。

1. 美国技术研究现状

美国在消费品安全信息分析、预警及控制等方面形成了较为完备的体系,主要通过以下几个方面来监测相关信息:

① 美国消费品安全委员会(简称 CPSC)。管辖范围的消费品约有 1.5 万多种,主要包括家用电器、玩具、儿童用品以及其他用于家庭、体育、娱乐及学校的消费品,其工作流程如图 1-1 所示。CPSC 最常用的三个信息系统分别是:国家电子伤害监测系统(National Electronic Injury Surveillance System, NEISS)、国际贸易数据系统(International Trade Data System, ITDS)和自动化商业环境(Automated Commercial Environment, ACE)。NEISS 收集来自美国境内急诊部门接触因产品造成的受伤病例信息,并可按照时间、产品种类、年龄段、诊断病种、事故场所以及受伤身体部位等,归纳与具体产品有关的急诊伤害案件总数;ITDS 与 ACE 协助 CPSC 掌握船货清单数据,并借此迅速找出具有危害的进口产品数据,以便实时处理。

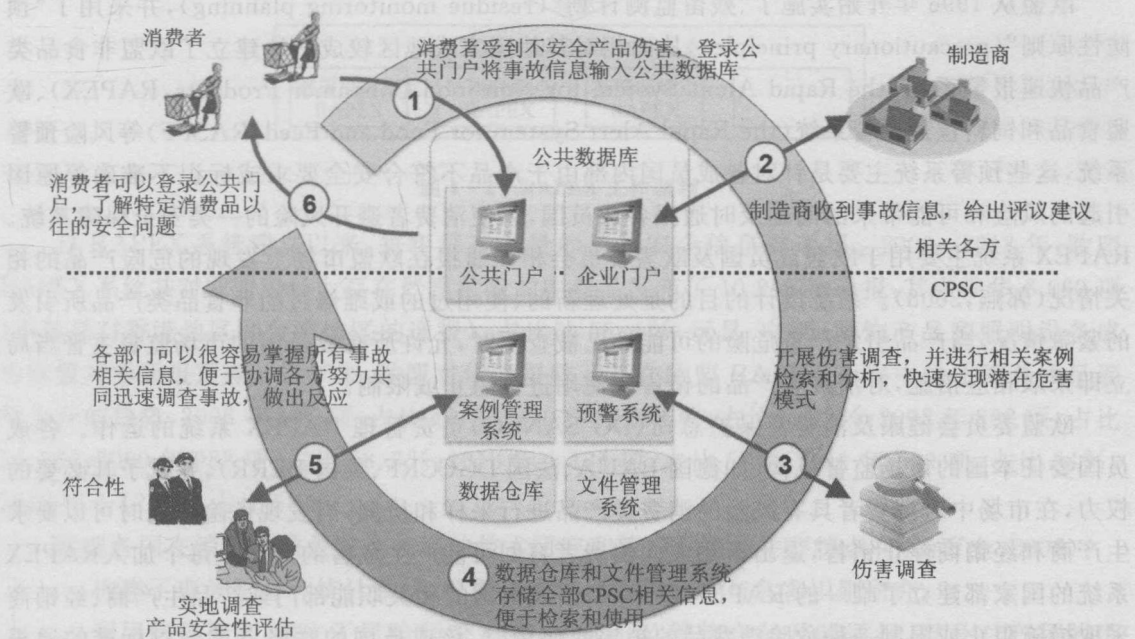


图 1-1 CPSC 工作流程

② 死亡证明文件(death certificate file)。如果发现消费品与死亡事故相关,则美国各州卫生部门会将这些死亡证明提供给 CPSC。交换中心将受害者信息移除之后,提供死亡证明



摘要。

③ 深入调查文件(IN-DePth investigations(INDP)file)。INDP 文件包含事件调查报告的摘要,调查围绕着与产品相关的伤害或事故。基于受害者/目击者的谈话,报告会提供事故序列、人的行为和产品的涉入程度。

④ 伤害/潜在伤害事故文件(Injury/Potential Injury Incident file(IPII))。IPII 文件包含了从热线报告、报纸报道的相关产品、体检医生的报告和 CPSC 收到的来信等渠道收集的信息的摘要,并按照消费品分类检索。

⑤ 1999 年 11 月,由劳工部所属的职业安全健康署(简称 OSHA)提出人因工程法草案法案(Ergonomics Program, Proposed Rule)。在该法案中规定了全套人因工程计划,其中包括:辨别工作中的人因危害,并对其进行分析与控制。美国许多大学都设有人因工程实验室,来研究产品在使用过程中的人因问题,为产品的使用安全提供了科学数据,同时也提供了一些消费品质量安全信息分析的技术方法。

美国在消费品质量安全领域的技术和管理方面的主要特点有:

- 建立了完善的消费品质量安全信息监测体系,并确立了覆盖比较全面的信息源;
- 充分利用了信息技术,提高了消费品质量安全信息获取、发布和快速应对能力;
- 重视消费品质量安全评估技术和产品使用中的人因分析研究,建立了消费品质量安全监控标准体系。

2. 欧盟技术研究现状

欧盟从 1996 年开始实施了“残留监测计划”(residue monitoring planning),并采用了“预防性原则”(precautionary principle),从而相对其他国家或地区较成熟地建立了欧盟非食品类产品快速报警系统(the Rapid Alert System for Non-food Consumer Products, RAPEX)、欧盟食品和饲料快速预警系统(the Rapid Alert System for Food and Feed, RASFF)等风险预警系统,这些预警系统主要是针对各成员国内部由于产品不符合安全要求或标识不准确等原因引起的风险和可能带来的问题及时通报各成员国,以使消费者避开风险的一类安全保障系统。RAPEX 系统主要用于欧盟成员国及欧盟委员会相互通报在欧盟市场上发现的危险产品的相关情况(郭燕,2008)。系统设计的目的是处理新的、使用过的或维修过的非食品类产品所引发的紧急情况,当产品引发严重危险的可能性已被查明时,允许所有成员国的市场监督主管当局立即采取相应措施,对消费者产品的销售或使用进行阻止或限制。

欧盟委员会健康及消费者保护总司(DG SANCO)负责管理 RAPEX 系统的运作。各成员国委任本国的市场监督机构(如德国 BAUA、法国 DGCCRF、英国 BERR),并赋予其必要的权力,在市场中对消费者具有风险的消费类产品进行采样和检测,当发现危险产品时可以要求生产商和经销商停止销售,退出市场或向消费者召回可能产生危害的产品。每个加入 RAPEX 系统的国家都建立了唯一的 RAPEX 联络点,当成员国的相关职能部门或产品生产商、经销商采取措施阻止或限制一种危险消费品的销售或使用,该成员国的联络点就会以标准的通报格式,向 DG SANCO 提交关于该产品的信息。标准的通报内容包括:产品信息、产品风险、防范措施、被通报产品的分销渠道等。DG SANCO 审查各成员国提交的信息,经过确认的通报信息将会被传播到 RAPEX 系统中。所有成员国的 RAPEX 联络点都会向本国的相关职能部门提供 RAPEX 系统中的共享信息。相关职能部门据此核查本国市场中是否有被通报的产



品,是否有必要采取适当的措施或行动。这些措施或行动通常包括禁止/停止销售、权威机构发布公共预警、产品撤架以及召回等方法。RAPEX 的工作流程如图 1-2 所示。

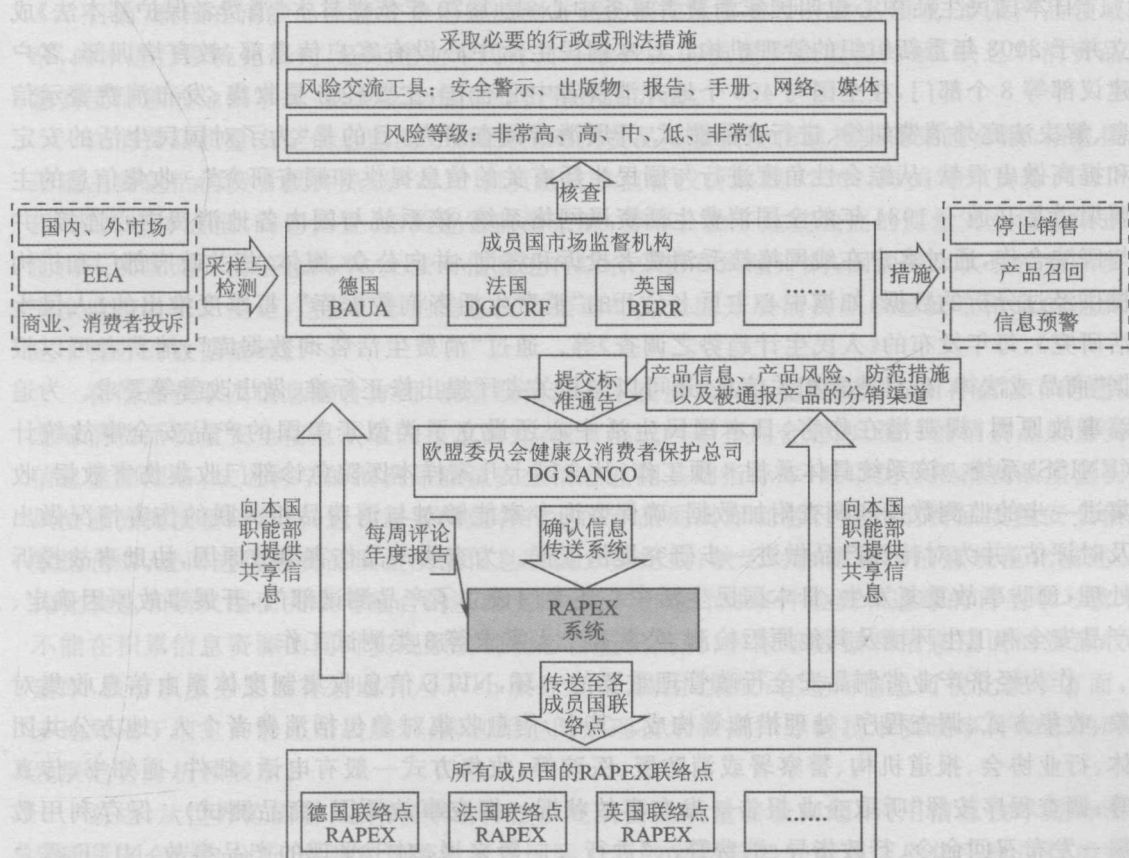


图 1-2 RAPEX 工作流程

自 RAPEX 系统建立以来,每年都有大量的产品安全信息被通报。2006—2012 年,欧盟 RAPEX 系统共对欧盟市场上存在严重危险的消费品发布了 10 946 项通报,其中涉华 6 060 项(不包括对香港地区和台湾地区的通报),占比达 55.4%,玩具、电器、服饰产品和照明设备成为欧盟对华通报重点。近年来,欧盟对华通报数及其在欧盟 RAPEX 通报中的比例总体呈逐年上升的趋势,2006 年 436 项,占比 47.8%;2007 年 683 项,占比 51.2%;2008 年 862 项,占比 54.8%;2009 年 982 项,占比 58.3%;2010 年 1 134 项,占比 58%;2011 年 839 项,占比 54%;2012 年 1 123 项,占比 58%。

欧盟各国在消费品质量安全领域的技术研究和管理工作方面的主要特点有(汤万金,2009):

- 构建了事故监督和统计系统,开展伤害事故原因分析与危害识别研究;
- 利用信息技术,建立了风险预警系统,实现了监管信息的及时通报,以及对产品的生产、经销、消费等不同环节的风险监控。

3. 日本技术研究现状

日本负责收集产品伤害信息并进行调查研究的机构主要有两个:日本国民生活中心