



企业竞争情报 智能挖掘

Intelligent mining for enterprise competitive
intelligence

张玉峰 李 纲 等著

学习出版社



企业竞争情报 智能挖掘

Intelligent mining for enterprise competitive
intelligence

主著者 张玉峰 李 纲 何 超
吴金红 王翠波
著 者 胡 凤 董坚峰 蔡皎洁
吴 覃 张 婧 周 磊
曾奕棠

学习出版社

图书在版编目 (CIP) 数据

企业竞争情报智能挖掘/张玉峰等著.

- 北京: 学习出版社, 2013. 7

(国家社科基金后期资助项目)

ISBN 978 - 7 - 5147 - 0329 - 0

I. ①企… II. ①张… III. ①企业竞争 - 竞争情报 - 情报检索 - 研究
IV. ①F274

中国版本图书馆 CIP 数据核字 (2013) 第 047722 号

企业竞争情报智能挖掘

QIYE JINGZHENG QINGBAO ZHINENG WAJUE

张玉峰 等著

责任编辑: 李 岩 张 俊

技术编辑: 王晓勇

封面设计: 杨 洪

出版发行: 学习出版社

北京市崇外大街 11 号新成文化大厦 B 座 11 层 (100062)

010 - 66063020 010 - 66061634

网 址: <http://www.xuexiph.cn>

经 销: 新华书店

印 刷: 北京市密东印刷有限公司

开 本: 710 毫米 × 1000 毫米 1/16

印 张: 22.5

字 数: 379 千字

版次印次: 2013 年 7 月第 1 版 2013 年 7 月第 1 次印刷

书 号: ISBN 978 - 7 - 5147 - 0329 - 0

定 价: 45.00 元

如有印装错误请与本社联系调换

国家社科基金后期资助项目

出版说明

后期资助项目是国家社科基金设立的一类重要项目，旨在鼓励广大社科研究者潜心治学，支持基础研究多出优秀成果。它是经过严格评审，从接近完成的科研成果中遴选立项的。为扩大后期资助项目的影响，更好地推动学术发展，促进成果转化，全国哲学社会科学规划办公室按照“统一设计、统一标识、统一版式、形成系列”的总体要求，组织出版国家社科基金后期资助项目成果。

全国哲学社会科学规划办公室

前 言

随着全球信息化和经济建设的快速发展，信息、知识、情报成为重要的战略资源，信息化成为企业和国民经济发展的重要工程，企业信息资源的深度开发与利用研究，成为学术领域和业界关注的热点。互联网、数据挖掘、人工智能等现代高新技术的应用与发展，为企业提供了海量信息资源开发与利用的良好条件，企业知识获取方法与技术的研究，成为全社会和企业竞争发展中的重要研究课题。

数据挖掘技术在企业、医学、金融等领域已得到广泛的应用，取得巨大的经济效益，并给企业和社会经济的各个方面带来了根本性的变革，同时，也为其进一步的深入研究提出了新的要求和更迫切的期待。数据挖掘的理论与应用研究成果，为企业竞争情报的挖掘与应用提供了新的思路 and 良好基础。融合多学科领域的先进理论与技术，革新竞争情报获取的理论和方法技术体系，使其向着自动化、智能化方向发展，势在必行，也必将目前信息层面的情报获取水平提升到知识层面的获取水平。

在这个跨越性的转变过程中，智能化的情报挖掘方法与技术是其关键枢纽。本著作综合应用情报学、数据挖掘、本体、人工智能、知识管理等领域的先进理论与方法，开创性地研究了企业竞争情报智能挖掘的基本理论、方法与技术，改进和完善了现代竞争情报获取方法技术体系，以实现从海量信息资源

中及时有效地挖掘情报知识，为企业管理决策提供智力支持。

《企业竞争情报智能挖掘》一书是在我们十多年研究及成果的基础上撰写的，也吸取了国内外相关学科领域的最新研究成果，较深入系统地研究了企业竞争情报智能挖掘模型、流程、策略、方法、技术及其应用，并注重理论与实践、创新及应用的结合。

全书共分为8章。首先讨论了数据挖掘的基本原理及其最新进展，分析了数据挖掘的原理和主要方法，探讨了数据挖掘向智能化发展的主要技术。重点从本体、Agent、知识情境、业务流程等方面构建了竞争情报智能挖掘模型；依据企业情报特征研究了企业竞争情报智能挖掘的基本流程——需求分析、信息搜集、信息整合、情报挖掘、结果评价；分别从信息内容、信息结构、信息行为、Deep Web、人际情报网络等角度深入探讨了企业竞争情报智能挖掘的实现策略、方法与算法；从语义层面研究和创新了企业竞争情报智能挖掘技术，包括企业竞争情报语义聚类挖掘、语义分类挖掘、语义关联挖掘、多维联机分析挖掘、融合本体和上下文知识的情报挖掘等。最后，将理论研究成果应用于实践，结合实例研究了金融领域的风险信息识别与挖掘方法和企业竞争中重要的客户知识挖掘方法。

全书由张玉峰、李纲主持撰写工作，负责全书的策划、统稿、定稿等，何超参加了统稿工作。参加撰稿的有（按撰写的章节顺序）：董坚峰、周磊撰写第1章；胡凤、曾奕棠撰写第2章；吴金红撰写第3、4章；王翠波撰写第5章；何超撰写第6章；吴覃、张婧撰写第7章；蔡皎洁撰写第8章。

在撰写本书过程中，我们引用和参考了国内外许多专家学者的论著，并得到了全国哲学社会科学规划办公室、武汉大学人文社会科学院、武汉大学信息管理学院、武汉大学信息

资源研究中心的大力支持，研究生唐涛、王志芳、徐斌、龙飞、黄恒、徐海峰、龚平、张苗苗、闻中慧、朱莹、孙超、吴宗朝、潘玉洁、钱景怡、丁雪做了大量资料收集与整理工作，在此一并表示最诚挚的谢意！

由于水平所限，加上竞争情报领域的研究内容复杂、难度较大，书中错漏之处在所难免，恳请读者批评指正。

本书系 2011 年国家社科基金后期资助项目（批准号：11FTQ002）的研究成果。

张玉峰于武汉大学

2012 年 9 月

目 录

第一章 知识经济时代的竞争情报	(1)
一、知识经济促进竞争情报的崛起	(1)
(一) 竞争情报的崛起	(2)
(二) 知识经济下的现代竞争情报	(9)
二、竞争情报研究的基本内容	(10)
(一) 竞争情报的功能和作用	(12)
(二) 竞争情报生命周期	(16)
(三) 竞争情报研究的重点领域	(18)
三、竞争情报采集	(21)
(一) 竞争情报采集现状及存在的问题	(21)
(二) 竞争情报采集的基本方法	(23)
(三) 竞争情报智能采集	(26)
四、竞争情报挖掘	(26)
(一) 竞争情报挖掘概述	(26)
(二) 数据挖掘技术在竞争情报挖掘中的应用	(27)
第二章 数据挖掘	(30)
一、数据挖掘概述	(30)
(一) 数据挖掘及研究现状	(30)
(二) 数据挖掘智能化进展	(32)
二、数据挖掘的主要方法与技术	(33)
(一) 概念分析	(34)

(二) 关联分析·····	(34)
(三) 分类分析·····	(39)
(四) 聚类分析·····	(42)
(五) 其他分析·····	(45)
三、数据挖掘智能化关键技术 ·····	(48)
(一) 文本挖掘·····	(48)
(二) 本体学习·····	(49)
(三) 语义分析·····	(62)
(四) 语义推理·····	(63)
(五) 智能检索与智能搜索·····	(66)
四、数据挖掘常用软件工具 ·····	(74)
(一) Enterprise Miner ·····	(75)
(二) Intelligent Miner ·····	(76)
(三) Clementine ·····	(77)
(四) MineSet ·····	(78)
(五) DBMiner ·····	(79)
 第三章 企业竞争情报挖掘模型 ·····	 (80)
一、企业竞争情报挖掘的基本模型 ·····	(80)
二、基于本体的语义挖掘模型 ·····	(82)
(一) 本体知识库·····	(83)
(二) 企业竞争情报领域本体的构建·····	(84)
(三) 语义知识挖掘·····	(88)
(四) 模型分析·····	(89)
三、基于 Agent 的分布挖掘模型 ·····	(90)
(一) Agent 技术 ·····	(90)
(二) 情报采集流程·····	(90)
(三) 智能挖掘模型·····	(91)
(四) 系统性能分析·····	(95)

四、基于知识情境的多维挖掘模型	(96)
(一) 知识情境	(96)
(二) 多维知识情境库	(98)
(三) 企业竞争情报多维挖掘策略	(98)
(四) 企业竞争情报多维挖掘模型	(100)
五、面向业务流程的挖掘模型	(101)
(一) 业务流程中的竞争情报	(101)
(二) 流程知识的获取方法	(102)
(三) 基于知识发现的流程挖掘模型	(104)
(四) 基于知识发现的流程挖掘系统	(110)
 第四章 企业竞争情报挖掘流程	(113)
一、竞争情报需求识别	(114)
(一) 竞争情报需求识别过程	(114)
(二) 传统需求识别方法	(115)
(三) 现代需求识别方法	(116)
(四) 竞争情报需求的表示方法	(117)
二、竞争情报收集与整合	(121)
(一) 企业竞争情报的信息源	(121)
(二) 数据选择	(124)
(三) 数据预处理	(125)
(四) 准备建模数据集	(133)
三、挖掘模型的构建	(136)
(一) 建立初始数据挖掘模型	(136)
(二) 检测初始数据挖掘模型	(136)
(三) 改进数据挖掘模型	(137)
(四) 评价数据挖掘模型	(137)
四、建模策略方法及实现	(139)
五、挖掘结果的处理	(140)

第五章 企业竞争情报挖掘策略	(142)
一、基于信息内容的挖掘策略	(142)
(一) 基于文本语义分类的信息内容挖掘	(143)
(二) 面向主题的信息内容挖掘	(144)
二、基于信息结构关联的挖掘策略	(147)
(一) Web 结构挖掘步骤	(148)
(二) 基于 Web 结构挖掘的动态竞争情报获取方法	(150)
三、基于信息行为的挖掘策略	(151)
(一) 面向信息发布者行为的挖掘策略	(152)
(二) 面向信息使用者行为的挖掘策略	(153)
四、基于 Deep Web 的挖掘策略	(153)
(一) Deep Web 信息采集中存在的障碍分析	(153)
(二) 面向 Deep Web 的知识搜索与提取	(158)
五、基于人际情报网络的挖掘策略	(160)
(一) 企业人际网络及其对竞争情报获取的作用	(161)
(二) 企业人际网络的挖掘流程	(164)
(三) 基于人际网络挖掘的竞争情报获取方法	(166)
第六章 企业竞争情报挖掘技术	(169)
一、企业竞争情报语义聚类挖掘	(169)
(一) 基于语义的聚类挖掘	(169)
(二) 企业竞争情报语义聚类挖掘算法	(171)
(三) 聚类挖掘实验与应用	(175)
二、企业竞争情报语义分类挖掘	(180)
(一) 基于语义的分类挖掘	(180)
(二) 企业竞争情报语义分类挖掘算法	(182)
(三) 分类挖掘实验与应用	(187)
三、企业竞争情报语义关联挖掘	(193)

(一) 基于语义的关联挖掘·····	(193)
(二) 企业竞争情报语义关联挖掘算法·····	(196)
(三) 关联挖掘实验与应用·····	(199)
四、企业竞争情报多维联机分析挖掘·····	(200)
(一) 基于语义的联机分析挖掘·····	(201)
(二) 企业竞争情报多维联机分析挖掘算法·····	(202)
(三) 多维联机分析挖掘实验与应用·····	(206)
五、融合本体和上下文知识的企业竞争情报挖掘·····	(210)
(一) 本体和上下文知识的融合实现·····	(211)
(二) 基于本体和上下文知识相融合的 竞争情报挖掘算法·····	(212)
(三) 融合本体和上下文知识的挖掘实验与应用·····	(216)
六、融合 Web 评论挖掘的企业竞争情报挖掘·····	(218)
(一) Web 评论挖掘·····	(218)
(二) 融合 Web 评论挖掘的企业竞争情报 挖掘流程·····	(219)
(三) 融合 Web 评论挖掘的企业竞争情报 挖掘算法·····	(222)
(四) 融合 Web 评论挖掘的企业竞争 情报挖掘实验与结果分析·····	(227)
 第七章 金融风险信息挖掘·····	(231)
一、金融风险概述·····	(231)
(一) 金融风险的定义及类型·····	(231)
(二) 金融风险的产生机理·····	(233)
(三) 金融风险分析挖掘的必要性·····	(235)
二、金融风险的识别方法·····	(235)
(一) 金融风险的定性分析·····	(235)
(二) 金融风险的定量分析·····	(237)

(三) 金融风险信息的挖掘方法	(239)
三、商业银行风险信息挖掘	(244)
(一) 商业银行风险概述	(244)
(二) 基于决策树的信用卡审批模型分析	(245)
(三) 基于粗糙集的欺诈风险分析	(248)
四、保险业风险信息挖掘	(252)
(一) 保险业风险概述	(252)
(二) 医疗保险数据分析实例	(256)
五、证券风险信息挖掘	(258)
(一) 证券业风险概述	(258)
(二) 证券市场信息分析实例	(260)
 第八章 客户知识挖掘	(265)
一、客户知识挖掘概述	(265)
(一) 客户知识挖掘的理论基础	(265)
(二) 客户知识挖掘的支撑技术	(266)
二、客户知识挖掘流程	(268)
(一) 客户知识挖掘目标分析	(268)
(二) 客户知识挖掘数据准备	(272)
(三) 客户知识挖掘模型建立及评价	(280)
三、客户知识的分析与预测	(299)
(一) 客户生命周期分析	(299)
(二) 客户价值分析	(302)
(三) 客户细分	(305)
(四) 客户响应预测	(307)
(五) 客户增值消费预测	(309)
(六) 客户流失预测	(311)
(七) 客户欺诈预测	(313)
四、商业银行客户知识挖掘实验	(314)

(一) 实验说明.....	(314)
(二) 客户知识挖掘流程.....	(315)
参考文献	(326)

Contents

Chapter One Competitive Intelligence in the Era of

Knowledge Economy	(1)
1. Knowledge economy has prompted the rise of competitive intelligence	(1)
(1) <i>The rise of competitive intelligence</i>	(2)
(2) <i>Modern competitive intelligence under the knowledge economy</i>	(9)
2. Basic content of competitive intelligence research	(10)
(1) <i>The role and function of competitive intelligence</i>	(12)
(2) <i>The life cycle of competitive intelligence</i>	(16)
(3) <i>The key field of competitive intelligence research</i>	(18)
3. Competitive intelligence acquisition	(21)
(1) <i>Acquisition status and existing problems of competitive intelligence</i>	(21)
(2) <i>Basic methods of competitive intelligence acquisition</i>	(23)
(3) <i>Intelligent acquisition of competitive intelligence</i>	(26)
4. Competitive intelligence mining	(26)
(1) <i>Overview of competitive intelligence mining</i>	(26)
(2) <i>The application of data mining in competitive intelligence mining</i>	(27)

Chapter Two Data Mining (30)

1. Overview of data mining	(30)
(1) <i>Data mining and its research status</i>	(30)
(2) <i>In intelligent data mining</i>	(32)
2. Main methods and technologies of data mining	(33)
(1) <i>Concept analysis</i>	(34)

(2) <i>Association analysis</i>	(34)
(3) <i>Classification analysis</i>	(39)
(4) <i>Clustering analysis</i>	(42)
(5) <i>Other analysis methods</i>	(45)
3. The key technology of data mining intelligent advances	(48)
(1) <i>Text mining</i>	(48)
(2) <i>Ontology learning</i>	(49)
(3) <i>Semantic analysis</i>	(62)
(4) <i>Semantic reasoning</i>	(63)
(5) <i>Intelligent retrieval and intelligent search</i>	(66)
4. Commonly used software tools in data mining	(74)
(1) <i>Enterprise Miner</i>	(75)
(2) <i>Intelligent Miner</i>	(76)
(3) <i>Clementine</i>	(77)
(4) <i>MineSet</i>	(78)
(5) <i>DBMiner</i>	(79)

Chapter Three Mining Model of Enterprise

Competitive Intelligence	(80)
1. Basic model of enterprise competitive intelligence mining	(80)
2. Semantic mining model based on ontology	(82)
(1) <i>Ontology knowledge base</i>	(83)
(2) <i>Domain ontology construction of enterprise competitive intelligence</i>	(84)
(3) <i>Semantic knowledge mining</i>	(88)
(4) <i>Model analysis</i>	(89)
3. Distribution mining model based on Agent	(90)
(1) <i>Agent technology</i>	(90)
(2) <i>Procedure of intelligence acquisition</i>	(90)
(3) <i>Intelligent mining model</i>	(91)
(4) <i>System performance analysis</i>	(95)

4. Multidimensional mining model based on	
knowledge context	(96)
(1) <i>Knowledge context</i>	(96)
(2) <i>Multidimensional knowledge context base</i>	(98)
(3) <i>Multidimensional mining strategy of enterprise</i>	
<i>competitive intelligence</i>	(98)
(4) <i>Multidimensional mining model of enterprise</i>	
<i>competitive intelligence</i>	(100)
5. Business procedure oriented mining model	(101)
(1) <i>Competitive intelligence in business procedure</i>	(101)
(2) <i>Acquisition method for procedure knowledge</i>	(102)
(3) <i>Procedure mining model based on knowledge</i>	
<i>discovery</i>	(104)
(4) <i>Procedure mining system based on</i>	
<i>knowledge discovery</i>	(110)

Chapter Four Mining Procedure of Enterprise

Competitive Intelligence	(113)
1. Competitive intelligence demand recognition	(114)
(1) <i>Competitive intelligence demand recognition process</i>	(114)
(2) <i>Traditional demand recognition method</i>	(115)
(3) <i>Modern demand recognition method</i>	(116)
(4) <i>Demand representation method of</i>	
<i>competitive intelligence</i>	(117)
2. Competitive intelligence acquisition and integration	(121)
(1) <i>Information source of enterprise competitive intelligence</i>	(121)
(2) <i>Data selection</i>	(124)
(3) <i>Data preprocessing</i>	(125)
(4) <i>Modeling dataset preparation</i>	(133)
3. Construction Mining model	(136)
(1) <i>Construction initialized data mining model</i>	(136)
(2) <i>Testing initialized data mining model</i>	(136)
(3) <i>Improving data mining model</i>	(137)