

内蒙古财经大学会计学术文库

基于大数据的企业财务预警 理论与方法研究

JIYU DASHUJU DE QIYE CAIWU YUJING LILUN YU FANGFA YANJIU

宋 彪/著



经济科学出版社
Economic Science Press

本书得到了国家自然科学基金项目“基于博弈论的信息安全理论与方法研究(编号:61272398)”、北京市社会科学基金重点项目“基于大数据的财务预警理论与方法研究(编号:14JGA001)”、内蒙古哲学社会科学基金“基于大数据的民族事件网络谣言识别与控制研究”的资助

内蒙古财经大学会计学术文库

基于大数据的企业财务 预警理论与方法研究

Research on Theory and Approach of Enterprise
Financial Early Warning Based on Big Data

宋彪著

经济科学出版社

图书在版编目 (CIP) 数据

基于大数据的企业财务预警理论与方法研究/宋彪著.
—北京: 经济科学出版社, 2015. 10
(内蒙古财经大学会计学术文库)
ISBN 978 - 7 - 5141 - 6182 - 3

I. ①基… II. ①宋… III. ①企业管理 - 财务管理 - 研究 IV. ①F275

中国版本图书馆 CIP 数据核字 (2015) 第 251409 号

责任编辑: 刘 茜 庞丽佳

责任校对: 靳玉环

版式设计: 齐 杰

责任印制: 邱 天

基于大数据的企业财务预警理论与方法研究

宋 彪 著

经济科学出版社出版、发行 新华书店经销

社址: 北京市海淀区阜成路甲 28 号 邮编: 100142

总编部电话: 010 - 88191217 发行部电话: 010 - 88191522

网址: [www. esp. com. cn](http://www.esp.com.cn)

电子邮件: [esp@esp. com. cn](mailto:esp@esp.com.cn)

天猫网店: 经济科学出版社旗舰店

网址: [http: //jjkxcs. tmall. com](http://jjkxcs.tmall.com)

北京万友印刷有限公司印装

880 × 1230 32 开 6.75 印张 200000 字

2015 年 10 月第 1 版 2015 年 10 月第 1 次印刷

ISBN 978 - 7 - 5141 - 6182 - 3 定价: 30.00 元

(图书出现印装问题, 本社负责调换。电话: 010 - 88191502)

(版权所有 侵权必究 举报电话: 010 - 88191586)

电子邮箱: [dbts@esp. com. cn](mailto:dbts@esp.com.cn)

| 前 言 |

随着经济全球化和互联网经济的发展，企业之间的竞争更加激烈，投资者和企业都更加重视财务危机预警在企业管理中的重要作用，然而企业财务危机预警一直是企业管理中面临的一个难题。当前的财务危机预警研究主要着眼点是基于财务指标建立模型，但是现实中每次经济危机发生时的预警悖论，使财务指标在危机风险预警方面的局限性逐渐暴露出来。财务指标的滞后性和可人为操纵性，严重地影响了财务危机预警的可信度。虽然有学者在模型中引入了非财务指标，但由于数据获取的难度，选取的非财务指标相对片面，难以适应对各种非财务指标具有不同敏感度的公司样本要求。大数据思维和技术为非财务指标的选择提供了新的思路。

本书研究基于大数据的企业财务预警的理论与方法。首先，针对财务危机预警这一难题，运用演化博弈论分析了财务危机预警准确度偏低的原因，证明了财务数据真实有效假设的不可靠性，发现了提高财务危机预警效果的有效路径。根据系统动力学仿真得出结论，指出发展相关信息获取技术是解决财务预警效果不佳问题的重要方向。在此基础上，提出了以网民为企业“传感器”的思想，基于互联网上的相关在线信息，通过情感分析处理，以及统计网民信息发布频次，融合后形成传感器信号，涵盖了与公司相关的所有维度的信息。利用基于大数据的网络爬虫和文本情感倾向分析技术，将与上市公司有关的信息结构化并且降维，形成关于上市公司的综合大数据情感指标，结合财务指标，最终建立引入大数据指标的财

务风险预警模型，并对模型的预测效果进行比较分析和检验，结果显示基于大数据的财务风险预警模型具有更好的有效性。相关结论从理论上提供了进行企业财务预警的新思路，也为相关方面预测上市公司财务危机提供了理论支持，最后指出作为 Web2.0 环境中在线信号的外在表现形式，企业网络舆情与企业危机是相互影响的关系，并根据“最好的预测是创造”的观点，从企业舆情的角度对企业危机管理进行了系统的研究。本书的选题来自于国家自然科学基金项目“基于博弈论的信息安全理论与方法研究（编号：61272398）”和北京市社会科学基金重点项目“基于大数据的财务预警理论与方法研究（编号：14JGA001）”的部分研究内容。本书的主要内容和创新点有以下几个方面：

1. 财务危机预警效果博弈论分析。针对企业财务危机预警的难题，用演化博弈论分析了财务危机预警活动的失败原因，论证了财务数据真实有效假设的不可靠性，指出了提高危机预警效果的有效路径。根据系统动力学仿真，明确了发展相关信息获取技术是解决财务预警效果不佳问题的重要方向。

2. 基于大数据的财务预警模型构建。将网民视为企业分布在互联网上的传感器，考虑到互联网上的相关在线信息，通过情感分析处理，获得能够涵盖相关公司的各个方面所有维度信息，通过结合财务指标，最终建立引入大数据指标的财务风险预警模型。

3. 大数据主题爬虫算法和语义分析算法设计。设计了可处理大数据的基于支持向量机和关键词的主题网络爬虫算法，构建了适用于财务领域的语义词典，进而设计了文本情感倾向分析算法，将互联网与上市公司有关的信息结构化和降维，收集具体传感器信号，支持本书设计的企业传感器模型，形成关于上市公司的综合大数据情感指标，进而对所设计模型的预测效果进行比较分析和检验，结果显示基于大数据的财务风险预警模型具有更好的有效性。

4. 基于大数据财务预警模型的企业危机管理。通过大数据财务预警模型可以发现，企业危机和企业舆情具有相互影响的关系。对

随时变化的企业舆情进行治理，是企业危机管理的重要环节。基于企业网络舆情在产生和发展过程中所呈现的群体性，应用群集动力学和演化博弈论的方法，在研究企业相关网络舆情群体流动过程和个体流动过程的基础上，构建了企业网络舆情疏导模型。通过使用多 Agent 仿真，应用所提出的模型，寻求在相关约束下的最佳疏导策略，证明了企业网络舆情疏导模型的有效性。研究成果为大数据财务预警模型提供预警信号的情况下，从网络舆情角度对企业进行危机管理提供理论依据。

目 录

第1章 绪论/1

- 1.1 选题背景和研究意义/3
- 1.2 主要研究内容/8
- 1.3 国内外相关研究综述/13
- 1.4 本书的创新之处/30
- 1.5 本书的结构/31

第2章 企业财务预警效果影响因素分析/35

- 2.1 财务危机预警效果影响因素/35
- 2.2 依靠监管提高预警效果的有效性分析/46
- 2.3 模型仿真与结果分析/55
- 2.4 本章小结/61

第3章 基于大数据的财务预警模型/63

- 3.1 问题的提出/64
- 3.2 理论假设/65
- 3.3 基于大数据的企业财务预警模型设计/68
- 3.4 模型仿真/84
- 3.5 本章小结/88

第4章 企业财务预警大数据指标数据获取算法 / 89

- 4.1 问题的提出 / 89
- 4.2 适用于大数据的 SVM 参数寻优策略 / 91
- 4.3 企业主题相关爬虫实现策略 / 93
- 4.4 算法实现 / 96
- 4.5 实验分析 / 98
- 4.6 本章小结 / 105

第5章 企业财务预警大数据指标数据处理 / 107

- 5.1 问题的提出 / 107
- 5.2 基于财经领域词典的企业财务大数据语义分析
总体算法 / 108
- 5.3 财经类网页爬虫算法 / 109
- 5.4 财经领域情感词典的建立 / 113
- 5.5 语义分析 / 115
- 5.6 实验分析 / 116
- 5.7 本章小结 / 119

第6章 基于大数据的财务预警模型实证分析 / 120

- 6.1 研究样本 / 120
- 6.2 研究方法 / 123
- 6.3 模型的构建 / 124
- 6.4 实证研究和结果分析 / 129
- 6.5 本章小结 / 136

第7章 基于大数据财务预警模型的企业危机管理 / 138

- 7.1 问题的提出 / 138
- 7.2 企业危机与企业网络舆情关系分析 / 141

- 7.3 基于群集动力的企业网络舆情演化博弈分析 / 149
- 7.4 演化博弈仿真 / 153
- 7.5 大数据技术背景下企业网络舆情疏导分析 / 156
- 7.6 基于大数据财务预警模型进行企业危机管理的
建议和措施 / 160

第8章 结论与展望 / 163

- 8.1 主要研究结论 / 163
- 8.2 研究局限与展望 / 165

附录 / 167

参考文献 / 189

后记 / 203

第 1 章

绪 论

近年来,随着经济全球化和社会信息化进程的加速,互联网和电子商务已成为引领社会经济发展的主导力量,互联网对企业的影响日益突出。与此同时,企业之间的竞争更加激烈。企业危机管理成为企业管理中的重要问题,而财务危机预警对于企业危机管理至关重要。

企业财务危机问题,是经营过程中关系到企业生死存亡的重要问题。企业财务陷入危机,轻则可影响企业的正常运转,损害信用机构、证券投资者等相关人的利益,严重时直接对国家的经济安全造成危害,甚至爆发世界经济危机。经济一体化的趋势日渐明显,电商的出现加剧了市场的瞬变性和难以预见性。企业间的竞争较以往更加激烈,2008 年金融海啸的爆发,使世界的经济产生了动荡,企业财务危机预警问题被提高到了前所未有的高度,受到了空前重视。财务危机风险预警是一个世界性的问题和难题^[1]。比较有影响力的财务预警方法诞生于 20 世纪 30 年代初期,到目前已经发展到十几种。每类方法的研究结果大都支持所采用的财务预警方法的预警效果,但验证于经济危机中的企业方面的证据极少。在具体实践应用中体现出,目前财务预警模型局限于事后确认的范畴,而很难达到事前发出警示的效果。究其原因,大多数模型中,财务指标是主要的预测依据。但财务指标往往只是财务发生危机的一种表现形式,甚至还有滞后反应性、不完全性和主观性^[2]。更为严重的是在基于财务指标预警模型建立过程中,学者们往往都假设财务数据是真实可靠的,然而这种假设无论在理论上,还是在现实环境中都有

待商榷。一些学者建立了加入非财务指标的模型,似乎可以给出正确的结论,当然,出于对财务指标进行修正的目的,加入非财务指标进行财务危机预警是必要的,这类研究在变量的选取上进行了积极的探索。但一般情况下,目前研究中所加入的能够起到作用的非财务指标都是依靠试错方法引入的,即都是在危机发生之后,才能够使某些指标得以确认以及引入模型,当下次危机类型不同,所建模型的局限性就会暴露出来。同时令人诟病的一点,引入的非财务指标大多都是考虑易获得性而非合理性,采用局部的数据进行验证,比如引入了如宏观经济变量、公司治理、产业变量和审计意见、市场份额、客户满意度等非财务指标。这样引入的非财务指标具有片面性,选取的过程也不够系统,无法顾及这些指标是否对所有企业具有适用性,也没有考虑非财务指标之间的组合作用,事实上不同的企业对一些非财务指标在财务风险方面的敏感度是有区别的。另外使用静态的非财务指标对具有动态属性的财务危机情形进行预测,忽略了事物发展过程中人的主观能动作用,颇有宿命论之嫌。种种这些原因都导致了在财务风险预警模型中,引入这些非财务指标变量的现实效果远不如研究中显著。

2015年1月,中国互联网信息中心(CNNIC)发布的第35次《中国互联网络发展状况统计报告》显示,截至2014年12月底,我国网民规模达6.49亿,互联网普及率为47.9%;全国企业使用计算机办公的比例为90.4%,使用互联网的比例为78.7%,企业互联网普及率已达到较高水平。随着各类互联网商业模式的发展,互联网与经济活动的深度结合,对传统商业模式的影响和改革程度将进一步扩大,互联网已经成为企业日常经营中不可分割的部分。因此,随着网络技术的发展,海量的个人网上搜索、发布或者关注行为,以及各种论坛、博客、媒体发布的企业相关信息往往在不经意间透露了真实的一些企业管理状态和走势信息,而且涵盖的范围非常的广泛。根据大数据的思维范式^[3],可否把这些信息,通过大数据的处理技术量化后引入财务风险预警模型当中?基于大数据的财务风险模型是否能够

在稳定性、有效性等方面具有一定的改善？这些问题的本质将是当前大数据技术破解财务风险预警难题的一个重要契机。

1.1 选题背景和研究意义

1.1.1 选题背景

一个企业的财务状况，是企业所有的利益相关人关注的焦点，这些利益相关人中包括经营者、企业债权人和投资者等。由于身处激烈的市场竞争环境中，风险已经成为任何一个企业无法回避的现实。如何进行企业财务危机预警，进行风险规避，是企业经营者进行市场经营决策，金融保险业的授信评级，以及投资者进行投资决策过程中必须考虑的重要问题。企业财务危机预警是一个悬而未解的难题，是一项在世界范围内进行了广泛而持续研究的热点课题。欧美学者研究企业财务危机预警较早，开始于20世纪中期左右。在20世纪末期，即90年代，在全球范围内资本市场得到空前发展，这导致了人们对财务风险进行管理控制需求的增加，学者们开始关注财务危机预警，并使之成为研究的热点。国内的资本市场形成较晚，市场经济体制的历史不长。证券市场尚无完备的监管机制，没有成熟的破产和退市规则，相对来说，国内学界对相关上市公司的财务风险危机的预测以及管控研究等起步处于落后阶段。普遍认同是20世纪末产生萌芽，在90年代随着中国的改革开放以及市场经济的飞速发展，财务风险预警研究吸引大量的研究者投身其中。由于在国内的市场环境中，国民经济中占有决定地位的是上市公司，而且由于市场机制的不成熟，只有上市公司才有可能提供完全符合信息披露制度标准的数据，方便信息的获取，使得国内的相关研究主要专注于上市公司。目前大多数研究者选择以财务指标构

建财务预警模型,即假设企业的实时经营运行状况能够被财务报表数据所真实的反映。但是通过 2011 年财政部进行调查之后,进而发布的上市企业会计信息质量的相关检查公告中,发现在财政部于 2010 年抽查的 114 家上市企业中,在企业内部控制、日常会计核算、合理缴税纳税这些方面有较大问题企业多达 24 家,达到了所抽查的企业总量的 21.05%。之后在 2012 年该公告指出仍有企业执行相关准则规范不到位,在会计核算等方面存在不同程度的问题。由于稽查技术的制约,现实中的会计失真,远比报告中披露的更为严重,因此在经济危机面前,大量看似财务健康的企业迅速地陷入困境^[1],这使一直以来的财务预警研究相关结果深受质疑。1997 年韩国的韩宝、三美、韩兴等大型集团企业,先后陷入破产倒闭的困境,结果导致金融机构急剧产生大量呆账贷款,使金融机构的金融资产状况,出现了严重困难,并且面临挤兑和破产等,由此诱发了韩国货币危机,进而发展为东南亚金融危机的严重爆发,最终影响了全球经济的稳定运行。在 10 年之后,即在 2007 年时,美国爆发了严重的次贷危机 (Subprime Crisis),引发了全球范围的金融市场震荡,直接导致大量的从事次级抵押贷款业务的企业崩溃,同时被迫关闭了大量的投资基金^[4]。雷曼兄弟公司于 2008 年 9 月在次贷危机中破产倒闭^[5]。随之而来的是希腊债务危机,爆发于 2009 年 12 月,同时陷入危机的国家囊括了葡萄牙、西班牙和比利时等,之后评级机构对意大利主权债务在 2011 年 11 月下调了评级,债务危机在欧洲的主要国家范围内已经呈蔓延的趋势^[6]。尤其值得特别注意的是——金融企业,虽然这些企业广泛运用了一些经典预测方法,比如 Altman 的多元判别分析等,进行了精细而全面的风险管理,但最终仍然在危机面前无能为力,在这次金融危机中成为重灾区^[1]。

理论上的成功与现实中失败的巨大落差,让企业对财务预警模型的相关研究感到了失望,甚至直接怀疑这些研究成果的有用性。基于财务报表建立的各种模型,由于会计信息的滞后和失真,无法实现财务危机预警的功能。根据现有文献可以发现,几乎所有研究

者在其研究过程中采取的预警方法都是可以通过论证的,相对在现实市场中的表现来说,这些预警方法是否有效,在现实企业环境中,往往找不到直接的证据。

在对财务危机预警问题的研究过程中,所构建的预警模型过于依赖财务指标,这使模型产生了一些先天的短板,为了能够补齐这些短板,有必要对传统的研究框架进行突破。我们必须对主要依赖财务指标建立模型的思路进行超越,应该发掘提高财务风险预警效果的路径,尽可能寻找导致企业危机的所有因素,以及发现影响企业危机预警效果的各种因素。在构建模型时,要考虑纳入这些因素,最终真正地使模型提高实际预警能力。一般认为非财务指标引入财务预警模型是提高预警准确度的有效途径,在对财务信息披露的相关经验研究中,由于学者技术的局限性,以及在研究过程中进行信息搜集和加工的困难程度,无论国外的学者还是国内的学者,几乎所有人的研究仅仅涉及那些比较正式的所谓书面披露,主要包括上市企业的年度报告和季度报告,以及国外像华尔街日报等,而在国内像三大证券报等披露的信息。或者进行一些访谈和调查,而对其他媒介披露的财务信息则几乎忽略不计^[7]。这种情况导致了引入的非财务指标具有很大的主观性和局限性,因而构建的包含非财务指标的预警模型表现也差强人意。

随着计算机和互联网技术在社会生活中的全面融入,信息的爆炸式增长已经达到了一个量变引发质变的程度。即信息总量的变化导致了信息形态的变化——量变引起质变。2008年《自然》推出名为“大数据”的专刊,指出“大数据”这个概念,最早产生于天文学和基因学,这些学科最先受到了信息爆炸的冲击。如今,几乎所有的研究领域中都有了“大数据”应用的影子。尤其 Web2.0 以来,网络信息飞速增长。在 2011 年 5 月的“云计算邂逅大数据”大会上,作为信息存储设备公司之一,即世界知名的企业 EMC 正式提及“大数据”概念。在同年 6 月,一些重要研究机构,像 IBM 和麦肯锡等随之提出了一些与“大数据”相关的成果汇报。这使“大数据”在计算机以及互联网行业迅速成为热点,如 IBM、Teradata、Informatica

和 NetApp 等世界闻名的 IT 企业都对“大数据”进行了研究投入。直观来讲,“big”表示信息数据量很大,所以在国内有学者认为它就是海量数据,或者认为是超大规模数据,很明显这种理解不能完全描述大数据的内在特征。其实在其他行业,像在天文学领域,环境生态学和生物学领域,以及电信和金融等一些领域一直都拥有着海量的信息数据,学者们从来没有将这些数据叫做“大数据”。其根源就是因为这些领域的信息数据量规模很大,但是这些数据的结构非常简单,传统的结构化技术足以支撑学者们对数据进行分析和处理的需求。但目前的情况是计算机尤其是云计算技术的出现,使得数据产生、传播、处理的费用大为减少,这促使了数据总量的飞速增长。据 IDC 公司的相关统计,截至 2010 年年底,世界总体信息数据量已经上升到 120 万 PB,预计到 2020 年,整个世界信息量仅指以电子为存储形式的数据,将增加到近 44 倍,达到 35ZB 之多,这样的规模在平均每两年几乎翻一番。

大数据技术的出现,使获得多维度的非财务数据成为可能,这为从非财务指标角度研究财务预警提供了新的契机。以大数据技术为支撑,把网民视为企业的传感器,对企业相关海量信息进行挖掘,可以发现作为企业传感器信号的网民在线信号波动与企业经营过程中财务风险之间的联系,进而引入大数据指标,进行财务危机预警模型的有效构建。

针对大数据在各个领域应用问题,研究人员进行了大量的研究,取得了很多成果。但总体来说,对大数据指标引入财务预警模型尚处于起步阶段,还有很多可以探索的余地和空间。本书基于博弈论的思想,以财务预警效果提高途径分析为主线,以信息传播的规律为基础,重点分析大数据引入财务预警模型的必要性和可行性,最终建立基于大数据的财务预警模型,进而指出企业网络舆情与企业危机的关系,提出了从企业网络舆情治理角度进行企业危机管理的有效措施。研究成果为基于大数据财务预警模型进行企业危机管理提供了理论模型与具体方法。

1.1.2 研究意义

财务预警作为一个集信息学、行为学、心理学、管理学、经济学理论于一体的复杂问题,在几乎所有社会科学领域被广泛研究,大数据技术的出现使财务预警的问题获得突破的契机,如何利用大数据技术提高财务预警效果,有效地提高企业的危机管理水平,具有重要的理论和现实意义。

1. 理论意义

对于财务预警效果不佳的问题,不仅要关注指标的选取和模型的构建,更要注重财务预警活动内部蕴涵的策略和方法,参与财务预警这一活动的群体,自身处于复杂的社会环境中,其间存在着复杂的博弈竞争关系,如果忽略了财务预警活动本身参与者的理性策略抉择,就会导致不断发展的模型一直无法最终解决预测效果不佳的难题。本书从博弈的角度分析财务预警效果提高的路径问题,寻求财务预警活动中的积极发展策略,从社会学的角度分析并构建了财务预警的相关模型和方法,这具有重要的理论意义。

提出了适用于大数据的财经主题爬虫算法,并构建财经领域语义词典,根据财经领域语义词典对企业相关在线信息进行语义分析,最终提出大数据在财务领域中的应用技术框架,同时也为大数据在其他领域的应用提供技术框架和理论参考。

充实了财务预警理论,大数据和财务预警问题都是学术界研究的热点,但大数据应用于财务预警问题目前还没有系统研究。大数据技术对海量非结构化数据的处理能力,决定了财务预警模型在引入非财务指标的过程中,指标属性方面会发生根本性变化。把分布在互联网上的网民看作是企业的传感器,根据传感器信号传回的大数据进行企业危机预警分析。通过深入研究相关大数据指标的属性及财务特征,构建基于大数据的财务预警模型和响应机制,提出了大数据在财经领域的应用思路,形成基于大数据的财务预警理论与

方法,这从信息学的角度丰富了企业财务预警理论。

对于企业危机管理问题,在基于大数据预警的基础上,根据与企业利益相关的网民群集活动的规律,分析企业在线传感器信号的外在表现形式(即企业网络舆情)与企业危机的相互关系,给出有效疏导模型,为大数据环境下企业危机管理提供了新的思路。

2. 现实意义

大数据概念诞生以来,就吸引了越来越多的研究者研究大数据在各个领域的应用。现在,有越来越多的企业认识到网络中的大数据存在巨大的市场价值。大数据和各个行业如何契合,如何使大数据的概念在具体的应用中得以落地,目前都没有具体的技术框架。本书开发出适用于大数据获取的财务主题爬虫算法,获取作为企业传感器的网民对相关企业的信号数据,并对数据进行财务领域的语义分析,从而提炼出相应的大数据指标,对财务领域涉及的大数据进行探讨应用,探索了大数据具体的应用途经,可以供其他领域大数据应用进行参考。

利用博弈论分析财务预警效果提高的途径,可以明确财务预警研究发展方向的重点。构建基于大数据的财务预警模型,提高财务预警模型的效果,有助于投资者及时获得企业财务危机的信息,进行准确的决策,同时企业管理者可以及时根据预警的结果,以企业在线传感器的实时信号为载体,对企业危机进行疏导治理,宏观上,能够使行业对整体经济发展状况作出评估,对经济危机的产生和发展做到未雨绸缪的重要参考作用。

1.2 主要研究内容

本书以上市公司财务预警为目标,主要进行了五个方面的研究:大数据技术背景下提高财务预警效果的影响因素分析、基于大数据的财务预警模型构建、大数据指标数据获取算法和数据处理算法设计、基于大数据的财务预警模型实证、基于大数据财务预警模