

电网企业风险导向内部审计系统总体框架设计研究^①

尹 心¹, 向 菲¹, 陈 涛¹, 王海强², 陶 丽², 李向华²

1. 国网重庆市电力公司电力科学研究院, 重庆 401123; 2. 西南大学 计算机与信息科学学院, 重庆 400715

摘要: 提出了一种电网企业风险导向的内部审计系统设计, 利用数据挖掘技术分析电网企业审计大数据, 识别电网企业可能存在的潜在风险因素, 对挖掘的风险因素进行量化评估, 为电网企业风险预防与建设规划提供决策性支持。

关 键 词: 电网企业; 审计; 数据挖掘; 风险评估; 决策支持

中图分类号: F239.1; TP399

文献标志码: A

文章编号: 1000-5471(2021)03-0073-05

电网企业是重要的国有能源企业。近年来, 在电网企业信息化水平提高的同时, 积累了大量业务数据^[1]。这些数据中蕴涵着潜在的知识和规则, 为企业内部审计分析提供了丰富的源材料。但是, 内部审计涉及到财务、工程、科技项目等多个领域, 需要审计的数据信息量庞大、信息格式多样。传统的人工审计模式已经越来越难以适应信息化发展的需求, 容易受到人力客观条件的限制, 造成审计结果出现偏差^[2]。采用计算机技术来辅助完成审计工作已经成为趋势, 具有良好的应用前景^[3]。

目前, 自然语言处理^[4]、分类数据^[5]等挖掘技术和机器学习模型已被广泛运用在审计领域。例如: 词频-逆文档频度(Term Frequency-Inverse Document Frequency, TF-IDF)技术能够提取审计文本的关键词, 为审计工作提供关注重点^[6]。Word2vec 技术可以有效地将文本转换成计算机可以识别的向量形式, 进而被计算机加以利用^[4]。数据挖掘技术在审计领域的运用, 可以推动电网企业的审计业务朝着智能化方向转型, 节省人力资源成本。

电网企业内部审计面临着因客观因素或者人为因素造成的审计人员错报或者检查风险, 因此风险分析在电网企业内部审计中起着重要作用。把风险导向审计的理论和方法运用到审计过程中, 以识别重点风险为出发点, 发掘电网企业运营的潜在风险因素, 有助提高企业风险管控能力。因此, 设计用于风险导向内部审计的计算机辅助系统具有较大的现实意义和应用价值。

1 风险导向内部审计系统总体框架设计

风险导向内部审计模式将风险理念引入到电网企业审计过程, 运用计算机完成智能化审计, 结合人工进行审计结果评估, 保证了审计结果的可靠性。在审计过程中, 风险导向内部审计模式以电网企业风险因素挖掘为导向^[8], 利用审计专业人员对电网企业风险管理、控制和治理等方面进行综合量化评估^[9], 根据量化结果对审计风险项目排序, 确定存有潜在风险的项目。

电网企业风险导向内部审计系统总体框架如图 1 所示, 通过数据挖掘来分析审计数据和生产、经营、

① 收稿日期: 2020-06-18

基金项目: 国家自然科学基金面上项目(61976181); 重庆市基础研究与前沿探索项目(cstc2018jcyjAX0274); 国网重庆市电力公司电力科学研究院科学技术项目资助(SGCQDK00NYJS2000111)。

作者简介: 尹 心, 硕士, 主要从事电力大数据分析及应用研究工作。

管理、维护等过程中潜在的风险，识别出风险事项，经过审计专业人员量化评估后，为电网企业公司提供风险管控的相关决策依据。

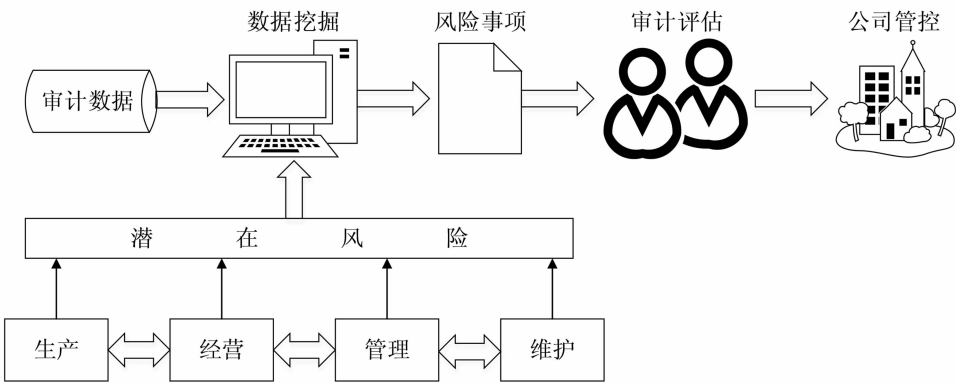


图 1 电网企业风险导向内部审计系统总体框架

2 风险导向内部审计系统技术构架

风险导向内部审计系统运用分类^[5]、离群检测^[1]、文本挖掘^[6]、自然语言处理^[4]等数据挖掘和机器学习方法实现审计数据智能分类与风险因素分析两大功能，其技术构架如图 2 所示。审计数据分类主要借助机器学习技术，首先将文本数据转化为向量，然后构建和训练文本分类器，完成对审计数据的分类。风险分析包含两个方面，即工程财务费用风险与企业风险管理关注风险。通过采用离群检测技术发现工程费用数据中存在的异常费用数据来发现工程费用财务风险，通过 TF-IDF 技术挖掘企业风险管理理论中的重点风险因素来发现企业风险管理关注风险，二者基于自然语言处理技术建立关联，进行电力企业风险挖掘。

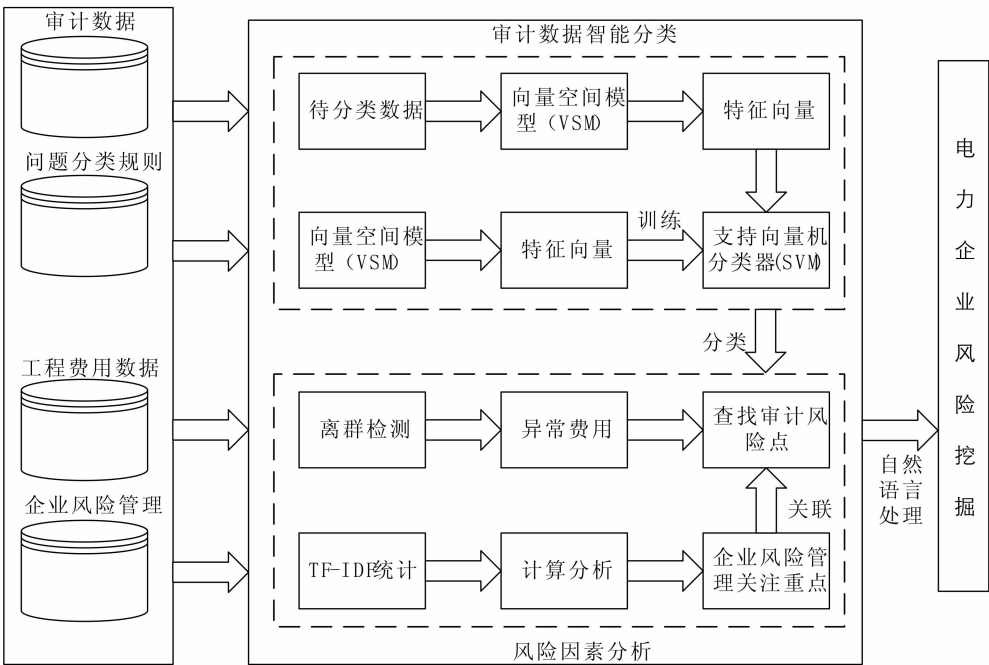


图 2 风险导向内部审计系统技术构架

2.1 审计数据分类

在审计领域，对发现问题的描述通常是以短文本形式呈现，这些非结构化的审计数据文本蕴含着丰富的审计信息^[5]。但是审计数据需要经过分类才方便进一步进行风险因素挖掘工作。采用人工模式分类整理海量的审计数据，结果不一定全面、客观、专业，容易受到审计人员自身能力的限制，而且消耗了大量的人力成本。本文设计了一种基于机器学习技术的审计数据智能分类方法，如图 3 所示。通过提取审计数据文

本数据特征并转换成向量形式, 然后输入到支持向量机(Support vector machine, SVM)分类器学习特征向量与文本类别之间的关系, 进而实现对审计数据的智能分类, 详细过程如下:

- 1) 提取文本特征向量: 采用向量空间模型(vector space model, VSM), 以每个词组为项, 词组出现频率作为项的权重, 将每个文本转换成向量 $V=(v_1, v_2, \cdots, v_n)$, 其中 v_i 表示第 i 个特征项的权重. 采用信息增益作为特征选择算法, 选取与分类相关的特征.
- 2) 分类器学习: 文本特征向量 V 及其类别标签 y 输入到分类器进行训练学习, 获得文本向量与类别对应的关系.
- 3) 审计问题分类: 待分类问题文本特征向量输入到 SVM 机器学习分类器, 实现对待分类问题的分类.

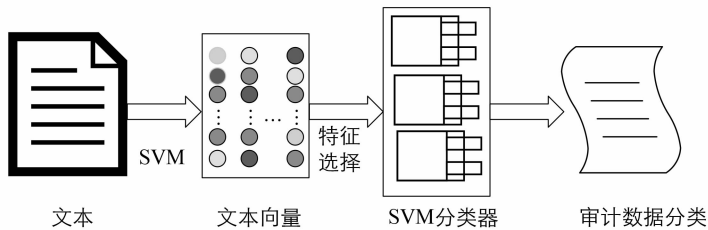


图 3 审计数据智能分类

2.2 风险分析

风险导向内部审计模式与企业风险管理具有相同的起点和目标, 即以风险识别和评估为起点, 以实现组织目的为目标^[7], 二者进行关联分析实现对风险的准确挖掘. 如图 4 所示, 风险导向内部审计模式用来挖掘财务报表中异常费用导致的风险, 并且计算分析企业风险管理关注风险项, 通过关联分析挖掘审计数据中的重大风险. 过程如下:

- 1) 异常费用检测: 运用离群检测技术^[1], 挖掘离群异常费用点;
- 2) 企业风险管理关注风险项识别: 运用 TF-IDF 挖掘分析企业风险管理理论关注的风险因素^[6];
- 3) 审计风险挖掘: 利用自然语言处理提取文本语义信息^[4], 建立异常费用问题描述与企业风险管理理论的关联, 挖掘出现费用异常的审计风险.

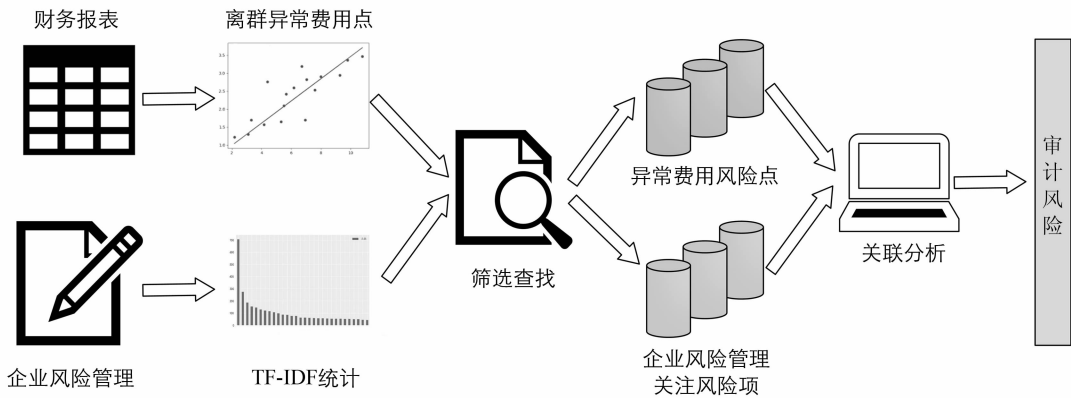


图 4 审计风险

异常费用风险点挖掘是通过分析财务及经营数据变动规律, 对成本、收入、资金流量、交易等过程中可能存在的异常财务及经营数据进行挖掘, 找出数据中与一般费用存在明显差异的异常数据. 在海量数据中, 离群检测技术可以有效地进行异常值的检测^[1], 在企业异常数据审计研究中已经证明该方法的可行性^[1]. 风险导向审计体系以电网企业积累的大量历史数据为支撑, 建立可靠的离群检测模型, 用来识别同类项目费用中明显偏离一般费用水平的异常项目费用. 对于出现异常费用的项目, 可以初步判定为存在潜在风险的项目, 作为风险因素挖掘的对象.

将异常费用风险点与企业风险管理关注风险项进行关联分析. 企业风险管理关注风险项主要采用 TF-IDF 统计方法对企业风险管理理论进行文本挖掘. TF-IDF 方法用来评估风险因素词汇在文档语料库中的重要程度, 可以及时过滤出现频率高但并不重要的词汇. TF-IDF 的值会随着词汇在文档中出现的次数的增加而增大, 也会随着词汇在语料库中出现次数的增多而减小^[6]. 通过 TF-IDF 统计分析计算后, 提取企业风险管理理论关注的风险因素, 以此作为存在异常费用的风险项目的风险评判理论依据.

根据审计问题分类, 查找含有异常费用的潜在风险项目对应的问题描述文本. 自然语言处理技术可以读取审计问题文本中的异常费用上下文信息^[4], 与企业风险管理理论关注的风险因素建立关联, 识别出异常费用对应的企业风险管理理论关注的风险因素, 建立异常数据变动与舞弊行为的关联关系, 自动化判断违规行为.

3 风险导向内部审计系统结果评估

风险导向内部审计系统所挖掘的审计风险结果需进一步结合专家知识和企业生产、经营、管理、维护等方面的实际情况进行评估, 并最终获得用于电网企业决策管理支持的风险因素清单. 风险导向内部审计系统使用者在获得由风险导向内部审计系统挖掘出的风险因素候选清单以后, 需会同领域专家, 结合电力企业内部现状, 从降低电网企业生产、经营、管理等过程可能造成的经济损失与社会危害出发, 依据“风险因素威胁性=风险因素标准分 $\times$ 经济损失 $\times$ 社会危害”的风险因素评分体系^[10]对风险事项进行量化. 在量化过程中, 可能重新涉及围绕风险导向审计模式挖掘风险因素的资料收集与调研工作, 根据各风险因素评分, 计算每个审计项目总风险值, 并且列出项目风险排序表, 作为风险导向内部审计的最终评估结果.

电网企业可根据风险因素的威胁性, 根据企业当前运营状况, 调整企业战略决策, 及时规避风险, 减少企业损失. 同时, 根据风险评估结果, 及时发现企业内部不足, 为企业调整提供一个明确方向. 从长远发展结果来看, 这将有助于转变传统审计中事后发现问题的弊端, 及早发现威胁企业发展的风险.

4 结 论

电网企业内部审计工作量大、涉及面广, 以风险为导向的内部审计模式能够在繁杂的审计工作中识别出重大风险事项, 为电网企业的风险管控决策提供支持. 本文以风险识别为出发点, 运用数据挖掘技术设计了以风险为导向的内部审计体系. 然而, 以风险为导向的内部审计模式还需要根据实际应用场景不断进行完善, 在复杂的审计数据和规则中, 利用信息技术准确识别出重大风险还有很长的路要走, 需要相关科研人员不断创新.

参考文献:

- [1] GUANG Z, WANG Y, WANG Q. Enterprise Financial Audit Modeling Research Based on Data Mining [C]//2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS). New York: IEEE Computer Society Press, 2016: 497-500.
- [2] YUYang, HOU Jing, LIHuan. Study on Continuous Internal Audit System Modeling and Application [C]//2019 International Conference on Artificial Intelligence and Advanced Manufacturing. Providence: ACM Press, 2019: 1-6.
- [3] JIANG L P. Research on the Application of Computer Aided Audit Technology [M]//Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2018: 921-927.
- [4] FISHER I E, GARNSEY M, HUGHES M E. Natural Language Processing in Accounting, Auditing and Finance: a Synthesis of the Literature with a Roadmap for Future Research [J]. Intelligent Systems in Accounting Finance & Management, 2016, 23(3): 157-214.
- [5] 伍 洋, 钟 鸣, 姜 艳, 等. 面向审计领域的短文本分类技术研究 [J]. 微电子学与计算机, 2015, 32(1): 5-10.
- [6] RAMOS J. Using TF-IDF to Determine Word Relevance in Document Queries [J]. Proceedings of the First Instructional Conference on Machine Learning, 2003, 242: 133-142.

[7] 章敏捷, 杨小秋, 周 勤. 风险导向工程审计体系建设浅探 [J]. 财会通讯, 2014(28): 93-95.

[8] 汪 啸. 基于风险评估的风险导向内部审计模式应用初探 [J]. 国际商务财会, 2012(5): 56-59.

[9] 杨临春. 风险导向型内部审计机制建立路径探析——基于中国华电集团公司的经验分析 [J]. 中国内部审计, 2015(8): 53-55.

[10] 赵东来. 电网企业风险导向内部审计信息化研究 [J]. 电力信息与通信技术, 2013, 11(8): 7-11.

On Design of Risk-Oriented Internal Audit System for Power Grid Companies

YIN Xin¹, XIANG Fei¹, CHEN Tao¹,
WANG Hai-qiang², TAO Li², LI Xiang-hua²

1. State Grid Chongqing Electric Power Research Institute, Chongqing 401123, China;
2. School of Computer and Information Science, Southwest University, Chongqing 400715, China

Abstract: It is a challenge to carry out internal audits manually with large amounts of data that is generated by power grid enterprises. Therefore, intelligent audit modelsthat incorporate several data mining methods are required to meet the challenges caused by big audits data. This paper proposes a system a risk-oriented internal audit system for power grid companies. The system uses several data mining techniques to analyze audit data,to identify potential risk factors, and quantitatively assess the identified factorsfor power grid companies. The system proposed in this paperis able to provide decision support for the risk prevention and business planning of power grid companies.

Key words: power grid companies; internal audits; risk-oriented internal audit system; data mining; deci-sionsupport

责任编辑 张 杓