

标识解析技术在电力行业招标风险识别中的应用

凤晓星

上海杰雍建设咨询有限公司，上海

收稿日期：2023年12月6日；录用日期：2024年3月6日；发布日期：2024年3月14日

摘要

工业互联网是一项连接全行业生产要素的泛用载体，依托标识解析的中间件技术，实现信息链对资源链、技术链和业务链的深度融合。电力行业的工业互联网二级节点于2022年底建成，快速形成了数据资源整合的使能要素，推动电力行业招标采购业务的智能化转型升级。风险识别作为招标业务的核心价值，对大数据技术赋能有着急迫的需求。本文围绕工业互联网标识解析技术在招标风险识别中的应用，介绍了招标多源异构数据挖掘和风险数字特征识别的新方法，以期为电子招标智能化发展相关领域提供新思路。

关键词

工业互联网，标识解析，电力行业，招标风险

The Application of Identification Resolution Technology to Bidding Risk in Power Industry

Xiaoxing Feng

Shanghai Jieyong Construction Consulting Co., Ltd., Shanghai

Received: Dec. 6th, 2023; accepted: Mar. 6th, 2024; published: Mar. 14th, 2024

Abstract

The industrial Internet is a universal container that links the productive factors of whole industry. Relying on the middleware of identification resolution technology, the information chain is deeply integrated with the resource chain, technique chain and operation chain. The second node of industrial Internet in power industry was built at the end of 2022. It forms active capacity of integration

with data resources in the short term and results in the intellectualized transformation and upgrading for bidding procurement operations. The risk identification, which acts as the key value for bidding operations, makes strong demands on big data technology. This article focuses on the application of industrial Internet identification resolution technology in bidding risk identification and introduces two new methods, which are multi-source heterogeneous data mining and risk digital characteristic identification, in order to provide new ideas for the intelligent development of electronic bidding.

Keywords

Industrial Internet, Identification Resolution, Power Industry, Bidding Risk

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

工业互联网是实现行业全生产要素体系性联通的泛用载体，通过数字赋能的全局协同、智能决策、边缘感知、敏捷响应和动态优化，推动行业信息链、资源链、技术链和业务链的深度融合，是新一代智慧工业的业态导向。标识解析是各类要素在工业互联网上互联互通的关键技术[1]，标识解析基于中间件技术，建设标识资源池、打破异构数据采集交互壁垒。

电力行业工业互联网体系建设，为招标采购管理提质增效提供了新的有利条件，标识解析技术应用用于招标大数据多态场景，重塑了数据挖掘、仓储、加工和利用的全链业务模式，提升了数据资源向数字资产的转化效率，在合规性分析、风险预警、数据资源利用、流程挖掘优化、业务成本管控等方面都有重要的价值体现。本文以工业互联网标识解析技术在招标多源异构数据挖掘、风险数字特征识别两处具体应用为例，以期对招标智能化相关领域提供创新思路。

2. 风控能力是招标业务的核心价值

招投标业务是一项具有社会化企业参与度高、依法合规性强的商业竞争活动，过程中存在专业跨度大、业务环节多、工作组织流程繁琐等现实情况。招标代理机构顺应市场经济发展需要，在专业能力和招标经验方面积累专业优势。受招标人委托为各类工程、货物和服务采购项目，提供合规比选、公平竞争的综合评价服务，在国民经济和社会发展上做出重要贡献。招标代理服务本质是一项商业交易法律化活动，招标代理机构作为提供招标代理服务的平台，依据国家及地方法律条款，依法执行招标程序，为采购人、投标人等参与方提供服务。高效优秀的招标风险防控能力，是招标代理机构体现服务能力、赢得客户信任的核心价值所在，也是招标代理机构打造品牌、立足市场、稳固经营的重要驱动力[2]。

随着电力行业招标采购方式、程序和标准的成熟，招标活动的市场竞争更加开放、灵活和多变，由此带来招标风险成因复杂、类型多样、特征隐蔽等问题[3]。招标代理机构作为招标活动的组织方、招标方案的制定方、招标过程的执行方，承担风险防控法定责任，需要通过不断提升科学管理水平从而保障业务核心价值的安全性、可靠性。当今电子招标成为行业主流趋势，风险识别的信息来源愈发丰富，招标管理创新、大数据建模、智能识别也已具备良好的实践条件。行业内对招标采购风控智能化的管理创新寄予强烈期望。

3. 工业互联网为风控智能化提供有利条件

国家工业和信息化部于 2020 年规划建设电力行业二级节点,以融合创新应用,实现产业链数字化协同转型作为发展方向。截止 2022 年底,行业标识注册量已过万,日解析请求达 5 万次,累计应用服务量 200 万次,服务企业涵盖设计、施工、监理、运维、物流、融资、保险、招标代理等。工业互联网因其强大的数据链接能力,为全过程智能招标风险控制,提供了数据资源挖掘上的便利条件[4]。非标识解析的传统模式下,各个应用端自行保管自有数据,并根据需求相互调取数据,交叉数据的质量和时效无法对齐。

在引入标识解析后,投标企业“人机财技信”风险要素形成了行业级统一标识链路,招标风险数据以企业为主体进行归集,数据管理转向“企信合一”模式,较之以平台为主体归集的传统模式,技术层面提供了数据多源异构的解决路径。另一方面,行业级数据域的可拓展性,突破了数据需求与应用场景的固化关系,招标风险数据的采集不再与风险识别模型绑定,数据需求从应用导向变为价值导向。

4. 基于标识解析多源异构的数据挖掘

招标风险智能化分析模型,包括数据溯源、因子识别、模型设计和效果修正四个部分[5]。在数据溯源方面,招标风险的基础数据存在多源异构、开放度不一、时效质量保障难的客观不利因素。招标风险数据源包括国家级和企业级两个数据域。国家级数据域以企业工商信息和信用信息数据库为主,诸如国家企业信用信息公示系统、信用中国、中国裁判文书网、中国执行信息公开网、全国认证认可信息公共服务平台、国家知识产权局政务服务平台。这些国家级数据库提供股权、资质认证、专利、公开征信、行政处罚等重要信息。公共数据源虽然长期开放访问,但对访问量有严格要求,单日访问量限制通常在 10~20 条,且有 3 分钟左右的应答延迟。对于项目体量较大的招标业务场景而言,业务链时效累积影响度估测在 3~5 天左右。不同数据源还存在相同对象的异构,如企业信用信息公示系统和信用中国平台都开放了行政处罚记录的查询,但二者覆盖口径略有差异,风险识别需要进行数据归集、格式转换、相似项合并。由于数据源访问能力的限制,多源头数据的时效偏差可控性不高,据估计电力行业中约有 87% 的招标项目,难以满足业务层面可接受水平[6]。企业级数据域的数据维度更加丰富,但仅在资格审核和开标两个时间点,允许作招标风险的数据调用,时效可控性上远不及国家级数据域。

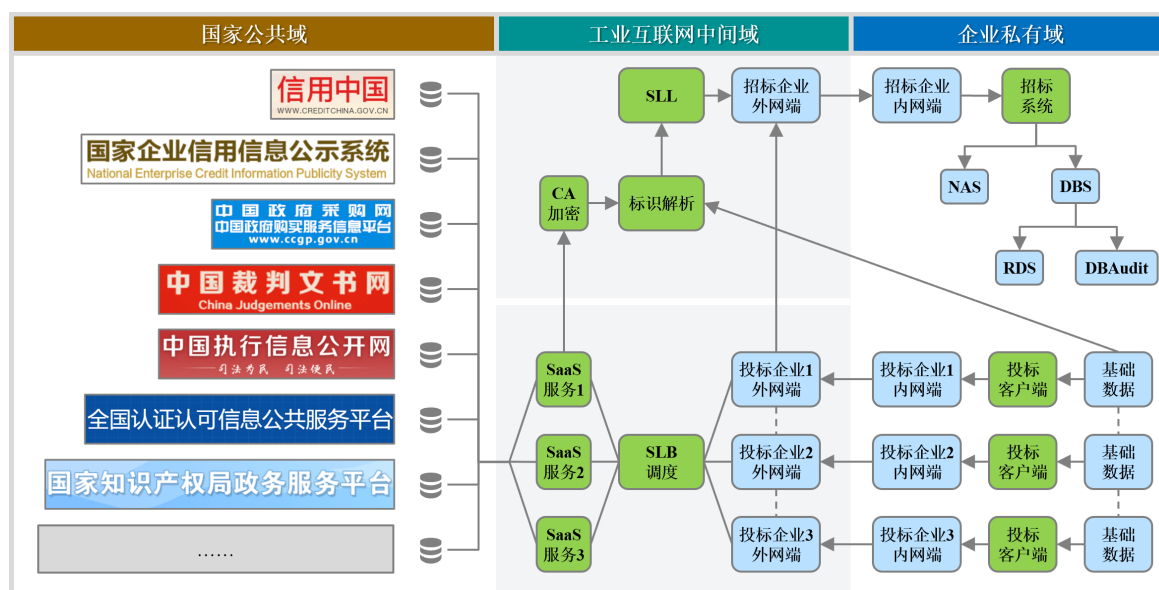


Figure 1. Bidding operations data structure based on the identification resolution intermediate domain

图 1. 基于标识解析中间域的招标业务数据架构

一种基于标识解析技术的解决方案是建立招标数据中间域，作为国家域和企业私有域中间的过渡域，较为简单的实现方法是部署 SaaS (Software as a Service)应用，实现公共域的分布式服务访问，化解单点访问量的冲击。采用 SaaS 应用模式的优势在于极强的可扩展性，企业可以根据自身需求进行数据管理和应用部署，而不用为了招投标业务单独进行云端管理，业务总体结构如图 1 所示。

SaaS 应用被部署于工业互联网云端，由 SLB (Server Load Balancing)负载均衡器负责请求调度。SLB 负载均衡器引入分布式服务器和虚拟 IP 地址，在遇到招标数据需求时，发起 SaaS 应用的远程数据访问功能。SLB 负载均衡器根据数据需求量，动态调整服务器资源，保证请求并发能力满足数据时效性要求。SaaS 模式向电力行业提供全过程的支撑，包括设计、生产、仓储、运输、采购、施工和检修等环节的应用部署。标识解析技术起到链接要素的作用，例如将施工设备的自动化控制系统，部分授权接入到 SaaS 应用中，或接驳项目工程师的档案数据库等其他数据。

作为招标风险识别的关键数据，企业公开征信数据被 SaaS 应用调取后，并不直接转发至招标企业的外网端口，这主要是出于数据安全层面的考虑，尤其是开标阶段的数据传输，根据招投标相关法律法规要求，电子招标不得以明文传输数据，进一步明确了数据防篡改的安全要求。为此通常采取 SLL (Secure Socket Layer)证书方式解决数据加密问题，最为常用的是招投标企业级 CA (Certificate Authority)证书[7]。招标风险数据自公共域获取后，被 SLL 应用直接进行 CA 加密，加密过程只需要用到公钥，而解密则必须使用私钥，并建立数据包的标识解析。加密数据包的传输方式类似于快递物流，数据包的网络地址被存储于标识解析服务器上，标识地址被招标服务器作了匿名处理，类似网购中的号码保护原理。第三方无从知晓数据包的真实网络地址，更无法拦截、篡改、伪造数据包。具体流程如图 2 所示。

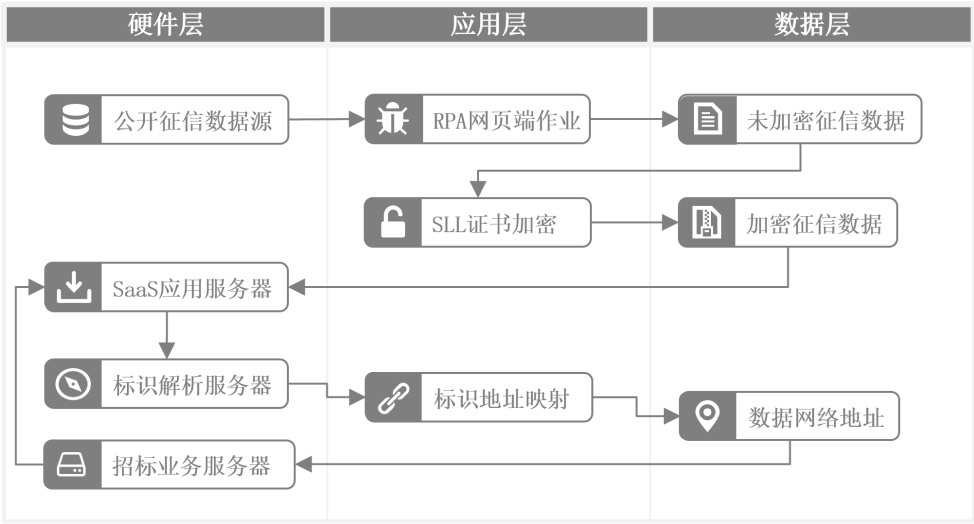


Figure 2. Encrypted transmission process for public credit data
图 2. 公开征信数据加密传输过程

该模式应用于多源异构数据处理，保留了源数据格式，数据转换标准和转换应用则部署于招标服务器上。一是，招标方根据自身风险分析的模型需求，自定义数据格式的转换，实现数据层、应用层和模型层的彻底分离，提高招标风险控制系统的适应性。二是，招标信用分析结果，反向传输回 SaaS 服务器，作为沉淀数据资产，实现招标大数据的资源归集效应。三是，信用数据除用于招标应用外，还可用于金融服务领域，为投标企业的融资贷款服务，进行信用风险的识别管控，实现数据链和价值链的重新组合。得益于标识解析技术，信用数据被定位到不同主体上，包括投标企业本体、法人股东、非法人股东、法定代表人和高级管理人员，招标和金融服务应用可以灵活穿透数据关系，数据挖掘过程受到的约束力极小。

5. 招标风险数字特征识别模型

招标风险类型十分宽泛，本文以法律风险和信用风险作为模型案例。法律风险来自招标活动中的关联交易，我国招标相关法律法规对公共采购中的关联交易作出了严格的限制性规定[8]，招标方应当设立事前审查程序，并于正式开标前完成审查，拒绝关联方进入后续环节。信用风险根据具体成因，常见的有投标方递交虚假投标文件、中标方拒绝签订合同两种[9]。

关联交易的风险识别，数字特征采用相关性分析，需要引入法人与自然人二级数据结构。法人层有招标方、投标方和招标(代理)机构，自然人层有组织的法定代表人和股东，以及评标委员会成员。数据范围包括主体名称、统一社会信用代码、法定代表人、身份证号码、股东列表、成立年限、注册资金。数据挖掘路径分为原始数据和解析数据，原始数据由投标方根据招标要求提供，解析数据部分由评标模块关键字智能抓取，另外一部分则取自第三方平台。为分析围串标的潜在风险，识别模型建立了项目维度的主体标识，项目主体数据被保存在标识解析中间域，和法人自然人一起构成三级主体关系。因此，围串标行为分析不再局限于单个项目，而是基于历史相关项目的整体数据情况，综合分析招标风险的程度，识别有效性较之前有大幅提升。

信用风险数据来源包括直接和间接两种形式。直接式信用信息是指原始数据已经对信用风险作了定性分类，如信用中国的失信被执行人，查询结果直接注明了查询对象是否具有信用风险。间接式信用信息指原始数据不注明信用风险大小，需要查询者进一步自行识别判断，例如中国裁判文书网的公开案件记录。直接式信用信息使用有限数集来表达风险分类，间接式信用信息必须依赖 NLP (Natural Language Processing, 自然语言处理)。为提高 NLP 执行效率，识别模型基于机器学习的原理构建。事前识别模型接受到一批风险企业的训练数据，风险企业被严格分为投标守信和失信两个类别。NLP 抓取信用信息的关键词，按照分布频率和语序配对关系，生成关键词关系图谱。识别模型读取训练数据后，经由关键词决策树建立二分法路径。关键词在决策树的分类下，被赋予正向和负向关联度，即通常意义下的褒义词和贬义词。识别模型不断迭代训练数据集，目标是建立最优决策树，使得语义分类可以最大程度与信用风险保持一致。语义分析风险识别的基本流程如图 3 所示。

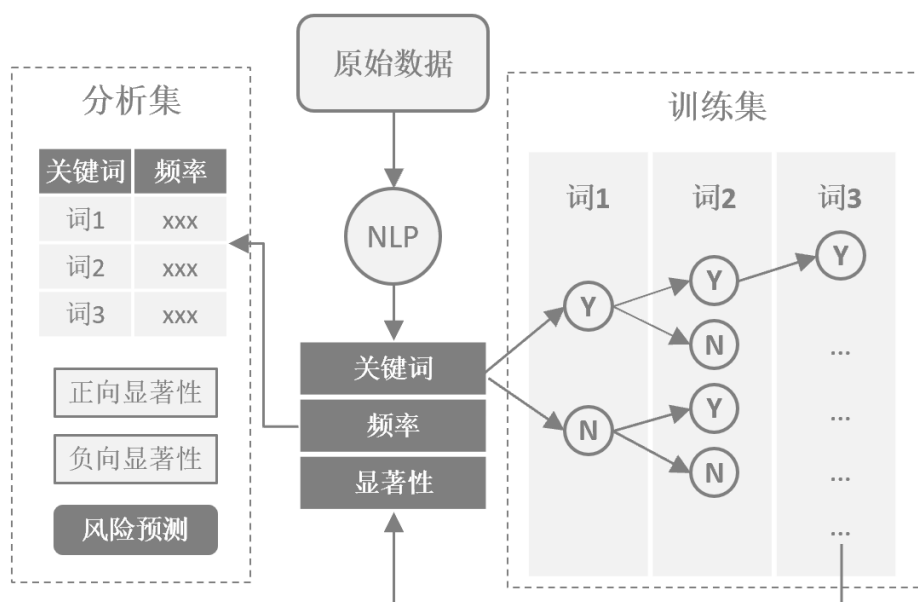


Figure 3. Identification of indirect credit risk based on LP significantly

图 3. 基于语义显著性分析的间接信用风险识别

6. 展望

工业互联网标识解析技术，在未来电力行业招标业务中，其主要价值在于完善行业级数据域，推动业务数据层和应用层的分离，形成项目主体的数据资产，辅助上下游企业共同参与包括招标在内的项目管理，并进一步引导标后成效数据进入招标风险模型中，提升全过程项目风险管理的闭环机制。

参考文献

- [1] 杜春燕, 呼彦朴, 徐华锦, 等. 工业互联网标识解析二级节点发展态势和展望[J]. 中国新通信, 2023, 25(4): 22-25.
- [2] 包雨霏, 蒋增平, 黄达春. 供应链管理视野下招标采购全面风险管理的探索[J]. 现代工业经济和信息化, 2023, 13(5): 251-254. <https://doi.org/10.16525/j.cnki.14-1362/n.2023.05.091>
- [3] 永学良. 电力物资招标采购风险分析[J]. 中国储运, 2023(5): 50-51. <https://doi.org/10.16301/j.cnki.cn12-1204/f.2023.05.088>
- [4] 罗开明, 王斌, 张崇见, 等. 电力装备行业工业互联网标识解析二级节点技术研究[J]. 电力信息与通信技术, 2021, 19(8): 47-53. <https://doi.org/10.16543/j.2095-641x.electric.power.ict.2021.08.007>
- [5] 彭茁, 王昭钦, 马红斌, 等. 智能化招标采购的合规风险防控[J]. 企业家, 2022(S1): 457-460.
- [6] 徐子雅. 电网公司招标采购管理优化路径研究[J]. 中国产经, 2023(8): 111-113.
- [7] 齐超然. 基于区块链的数字证书管理机制研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工程大学, 2021. <https://doi.org/10.27060/d.cnki.ghbcu.2021.000762>
- [8] 徐芳芳. 如何实现对公共资源交易中“关联交易”行为的法制化监管[J]. 法制博览, 2020(12): 186-187.
- [9] 邓晨曦. 电力行业招投标活动中供应商的信用管理探讨[J]. 电气技术与经济, 2022(5): 187-190.