

基于隐私计算的油气管网数据安全共享技术平台研究与应用

任武¹, 张珂¹, 富影杰², 张正雄¹, 蔡永军¹

1. 国家石油天然气管网集团有限公司科学技术研究总院分公司, 天津 300450;

2. 国家管网东北分公司沈阳油气计量中心, 辽宁 沈阳 110000

摘要

为保障油气管网关键信息数据资源安全性, 进一步释放数据价值, 对基于隐私计算的油气管网数据安全共享技术平台进行研究与应用。首先, 从政策、技术、行业三方面进行油气管网数据安全共享现状分析, 揭露数据安全合规与数据共享需求之间的矛盾; 然后, 针对这一问题设计实现油气管网数据安全共享平台, 并对3项关键算法原理进行详细描述, 实现“原始数据不出域, 可用不可见”和“可控可计量”; 最后, 针对油气管网行业数据安全共享场景进行梳理, 并结合3个典型应用场景进行数据安全共享技术应用。结果表明, 基于隐私计算的油气管网数据安全共享技术平台有效实现了油气管网企业内部数据交换、外部数据共享, 打破了数据之间的壁垒, 保障了数据在各业务链间的安全共享流通, 提升了油气管网管控水平。

关键词

隐私计算; 数据安全; 数据共享; 油气管网

中图分类号: TP309.2

文献标志码: A

doi:10.11959/j.issn.2096-0271.2024072

Research and application on data security sharing technology platform for oil and gas pipeline network based on privacy computing

REN Wu¹, ZHANG Ke¹, FU Yingjie², ZHANG Zhengxiong¹, CAI Yongjun¹

1. PipeChina Institute of Science and Technology, Tianjin 300450, China

2. Shenyang Oil&Gas Measurement Center, PipeChina Northeast Pipeline Co., Ltd., Shenyang 110000, China

Abstract

To effectively ensure the security of key information and data resources in the oil and gas pipeline network, and further unleash the value of data, research and application of a privacy computing based oil and gas pipeline network data security sharing technology platform are conducted. Firstly, analyze the current situation of data security sharing in oil and gas pipeline networks from three aspects: policy, technology, and industry, and expose the contradiction between data security compliance and data sharing needs. Then, design and implement a data security sharing platform for oil and gas pipeline networks to address this issue, and provide detailed descriptions of three key algorithm principles, achieving "original data not out of domain, usable but invisible" and "controllable and measurable". Afterwards, the

data security sharing scenarios for the oil and gas pipeline network industry will be sorted out, and data security sharing technology will be applied in combination with three typical application scenarios. The results indicate that the data security sharing technology platform for oil and gas pipeline networks based on privacy computing effectively realizes internal data exchange and external data sharing of oil and gas pipeline network enterprises, breaks down barriers between data, ensures secure sharing and circulation of data between various business chains, and improves the level of oil and gas pipeline network management and control.

Key words
privacy computing, data security, data sharing, oil and gas pipeline network

0 引言

能源是推动现代工业化进程的核心和驱动力，作为油气资源传输的关键途径，油气管网已成为支撑国民经济发展的关键基础设施^[1-4]。国家能源局 2023 年提出促进油气管网进行信息化和数字化提升的多项建议，旨在推动智能管道、智能化储油罐和储气库的发展，从而提高油气管网设施的安全性、高效性、可靠性^[5-8]。在数字化转型、智能化应用的背景下，各种数据源大量涌现，油气管道企业也适时地提出建设“油气管道数据‘全国一张网’”的战略发展目标，为智慧管网的构建提供数据驱动和平台支撑，打通企业内网与互联网的数据通道，为国内外管道企业提供开发的数据资源、数据标准。

随着数据量激增和数据潜在价值的持续提升，数据已成为推动油气管网发展的核心资产。然而，数据时代的到来不仅带来了数据价值的飞跃，同时也伴随着数据安全事件的频繁发生，这些安全事件所造成的危害和损失不亚于生产安全事件。更为严峻的是，当前数据要素在安全流通与共享的过程中遭遇了多重挑战。一方面，现有体系尚缺乏一种能够平衡数据安全合规与数据共享需求的合作机制与技术路径。这种缺失导致数据主体之间在秘密泄露、

利益分配等方面存在难以弥合的信任差距^[9]，这种信任缺失不仅阻碍了数据的有效流通与共享，还增加了数据保护的成本和技术门槛。另一方面，黑灰产业链的存在进一步加剧了数据共享的风险。不法分子通过非法手段窃取、贩卖数据，不仅侵犯了数据主体的合法权益，还严重干扰了正常的数据流通秩序。以 2021 年美国科洛尼尔管道运输公司遭受的网络攻击^[10]为例，黑客通过非法手段窃取了近 100 GB 的数据，导致该公司被迫切断网络连接，超过 8 850 km 长的管道系统关闭，全国能源供应中断，油价波动，社会不安。这一事件充分暴露了数据安全问题的严重性和紧迫性。

针对以上问题，本文对基于隐私计算的油气管网数据安全共享技术平台进行研究，对平台的功能架构进行详细设计，并系统性地阐述了数据安全共享技术的核心算法，之后结合油气管网 3 个场景进行技术应用，缓解了数据安全合规与数据共享需求的矛盾，打破油气管网行业的“数据孤岛”，促进数据安全共享与流通，进一步挖掘数据价值，探索开发新技术新产品。

1 现状分析

1.1 政策现状

自 2008 年以来，数据要素流动对全球

经济增长的贡献已经超过传统的跨国贸易和投资，成为驱动全球经济增长的最重要驱动力。2017年12月，习近平总书记在中共中央政治局就实施国家大数据战略进行的第二次集体学习中强调，“在互联网经济时代，数据是新的生产要素，是基础性资源和战略性资源，也是重要生产力，要构建以数据为关键要素的数字经济”。美国是最早对开放数据进行科学管理与共享的国家，美国早在1996年的《健康保险流通与责任法案》就开始关注隐私和数据安全，多集中在各州政府层面。2022年美国众议院和参议院联合发布了《美国数据隐私和保护法案》讨论稿，这是首个获得两党两院支持的全面的联邦隐私立法草案。在2021年《数据保护和隐私中网络安全措施的技术分析》中，ENISA建议各机构进行常规的“安全和数据保护风险评估”，以确定是否需要通过加密隐私增强协议降低数据处理中的风险。

我国数据要素相关法律法规等政策性文件也相继出台，2023年工业和信息化部等十六部门在《关于促进数据安全产业发展的指导意见》中提到“优化升级数据识别、分类分级、数据脱敏、数据权限管理等共性基础技术，加强隐私计算、数据流转分析等关键技术攻关”，对推动数据安全高质量发展，保障各行业各领域数据安全具有重要意义。同年，国家能源局在《关于加快推进能源数字化智能化发展的若干意见》中提到“深入实施创新驱动发展战略，推动数字技术与能源产业发展深度融合，加强传统能源与数字化智能化技术相融合的新型基础设施建设，释放能源数据要素价值潜力”，这对于激活数据要素潜能、带动能源网络各环节的互联互通具有重要意义。2024年1月国家数据局发布《“数据要素×”三年行动计划（2024—

2026年）》，旨在充分发挥数据要素乘数效应，赋能经济社会发展。

数据要素的合规高效流通使用是我国数据基础制度建设、促进要素价值发挥的核心任务，而数据安全共享是数据要素流通使用、充分释放数据要素价值的前提保障。随着各类制度文件的出台，2024年我国数据要素市场建设探索进入了关键时期。

1.2 行业现状

全球及国内管道行业正步入数字化转型的关键时期，顶尖油气企业已率先将数字技术渗透至油气产业链的每一个层面。例如，沙特阿美公司应用数字化、物联网、大数据技术推动传统油气产业转型升级，实现运营实时可视化、资产预测性维护、资产管理精细化。由于物联设备、数据采集设备和内置传感器等越来越多，油气运营数据在1年内预计增长超过16%，甚至达到30%。IDC调研显示大多数石油与天然气公司计划在未来2~3年会增加对数据管理与人工智能技术的投资。可见“数字化管道”作为新兴方向，正为油气管道工程建设与运维引入一种创新途径与手段。

目前，中国油气管网基本实现自动化，正在向智能化方向发展，智能管网以数字化管理为主要特征，在自动化感知与数据共享基础上，基于规则与流程将实时的数据、多维的规则等按照流程推送给设备或管理者指导决策作出响应^[11]。油气管网系统的关键业务数据，即SCADA数据，构成了调控操作的中枢，伴随着自动化与智能化的进步，不仅SCADA数据在核心层面得到持续扩展，而且各类检测和监测数据也在不断积累，这些数据增强了管道运营的管理效率和能力^[12]。然而，尽管油气管道行业内部积累了海量的数据资源，其

在社会经济发展与民众日常生活中的应用潜力尚未得到充分挖掘与释放。经深入分析,此现状主要归因于以下3点。

(1) 数据碎片化与分散化挑战显著。系统间、部门间、行业间建设标准与数据规范缺乏统一,导致业务互操作性差、数据流通障碍重重,形成了众多的“信息孤岛”,阻碍了数据的有效整合与协同利用,限制了信息价值的最大化发挥。

(2) 数据合规性与安全监管严苛要求。鉴于管道行业数据往往涉及国家安全与行业敏感信息,其保密性要求极为严格。当前,由于缺乏跨域数据融合计算的有效安全保障机制,数据安全管控措施也趋于保守,数据跨域共享与融合的实现面临重重困难。

(3) 数据融合应用能力不足。受限于跨域数据融合障碍,基于多源数据的人工智能与机器学习应用难以开展,严重制约了大数据技术在管道行业中的深度应用与落地。

1.3 技术现状

根据《关键信息基础设施安全保护条例》,油气管网属于国家关键信息基础设施,产生的数据涉及社会民生,具有高度敏感性,一旦遭到破坏、丧失功能或数据泄露,可能严重危害国家安全、国计民生、公共利益。近年来针对油气管网数据的攻击时有发生,在内部数据交换、外部数据共享以及与其他行业数据进行交互应用时,需对数据的安全性做到最大限度的保障。

国内方面,由于数字化时代的到来,各行业利用海量传感器、智能设备收集大量数据,并利用大数据技术进行数据分析和潜在价值挖掘。而传统的隐私保护技术

在面对这种规模大、来源多、动态更新等特点的数据时已不再适用^[13]。在技术创新上,国内企业在数据安全共享领域取得了显著进展。例如,通过采用联邦学习、隐私计算、可信任计算环境、差分隐私、多方安全计算、区块链等技术,实现数据的安全共享和分析^[14];通过联邦学习技术,打破行业间数据的孤岛^[15];通过构建零信任安全架构,提升数据安全共享的防御纵深^[16];通过区块链技术,规避集中式服务器数据泄露的风险^[17]。此外,国内企业还积极探索数据安全共享技术在油气管网领域的应用,推动了数据安全共享技术的落地实施。

国外方面,Gartner 机构认为,数据安全治理远非单一工具集或产品级解决方案,而是一个自组织决策层向下延伸至技术执行层,横跨管理制度与技术支持双重维度的系统性工程^[18]。这一治理体系自上而下地贯穿于整个组织架构之中,形成了一条紧密衔接、协同作用的完整链条。Gartner 在其 2023 年的数据安全成熟度曲线中新增了 5 项新兴技术,包括加密敏捷性、后量子密码学、量子密钥分发、主权数据策略和数字通信治理。这些技术旨在应对量子计算威胁、安全工具的融合和集成,以及管理未知的影子 IT 数据。针对数据安全领域的挑战,微软作为全球科技巨头推出了一系列数据安全治理方案。该方案植根于微软自主研发的隐私、保密及合规性框架之内,旨在通过一种统一的、跨学科的策略方法,实现对数据边界安全性的有效维护。此数据治理计划不仅强调技术的先进性,更注重策略与实践的深度融合,确保了数据安全措施在复杂多变的业务环境中能够持续发挥效用。

虽然国内外企业在数据安全共享技术领域取得了一定进展,但是仍然面临一些

技术挑战。一方面,数据安全合规是当前企业面临的重要挑战之一,如何在确保数据安全合规的前提下实现数据共享,是一个亟待解决的问题。目前,虽然有一些技术模式(如区块链、隐私计算等)可以在一定程度上解决数据安全和隐私保护问题,但如何将这些技术模式应用于实际的数据共享场景中,并建立起一套完善的合作机制,仍然是一个技术难题。另一方面,数据主体之间在秘密泄露、利益分配等方面存在的信任差距难以弥合。数据共享涉及多个数据主体的利益,如何在确保数据安全的前提下实现利益的合理分配,是一个需要解决的问题。同时,由于数据泄露等安全事件时有发生,数据主体对数据共享的安全性存在担忧,这也增加了建立信任机制的难度。

因此,本文提出基于隐私计算的油气管网数据安全共享技术平台研究与应用,通过集成先进的隐私计算技术,建立一套完善的数据安全共享合作机制,并设计实现油气管网数据安全共享平台。其不仅能够缓解企业数据合规与数据安全共享的矛盾,还能有效实现数据“可用不可见”,避免数据在分析计算中被泄露,促进油气管网行业数据资源的深度挖掘和应用。

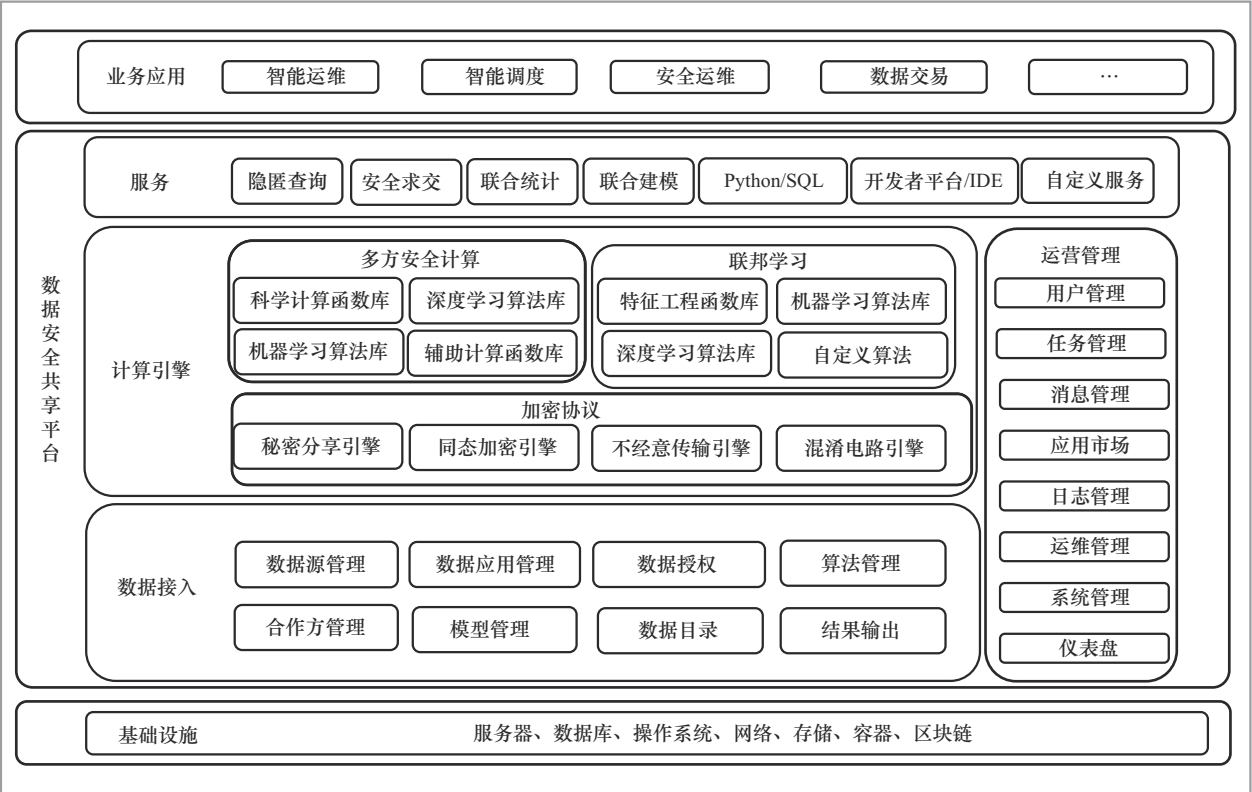
2 油气管网数据安全共享平台

针对以上问题,本文在剖析数据安全共享的痛点问题的基础上,设计实现油气管网数据安全共享平台。该平台采用了安全求交、隐匿查询、联邦学习等技术,确保数据可以在加密状态下进行计算和分析,保证数据共享流通过程中的安全性。本节将详细阐述该平台的架构设计及其关键算法原理。

2.1 平台架构

本文设计油气管网数据安全共享平台,其功能架构如图1所示。平台由六大功能模块组成,分别是基础设施模块、数据接入模块、计算引擎模块、运营管理模块、服务模块、业务应用模块。基础设施模块基于企业现有资源提供计算、存储、网络、区块链等基础设施服务。数据接入模块负责连接用户数据源与计算服务,保障加密数据传输至计算集群并解密结果。计算引擎模块由多个网络互联的计算节点组成,节点间通过安全通信互联,提供密文计算接口,并按多方安全协议执行密文计算。运营管理模块负责整体计算资源的管理、计算任务管理和调度功能,驱动数据接入和计算引擎进行相应工作。服务模块从多个业务中抽象出场景,包括基于多方数据的基础计算、联合统计、隐匿查询、隐私求交、联邦学习、SQL查询等。业务应用模块将企业内部以及外部业务合作方的多种类型数据进行安全融合,提供联合风控,联合营销、集团监管等业务。

该平台基于多方安全计算等技术实现既分享数据,又保证分享的数据明文不出本企业的目的,从而保证数据的“原始数据不出域,可用不可见”,并且规定数据的用途和用量“可控可计量”,解决了数据共享融合与数据安全合规这组核心矛盾。平台的优势在于算法、算力和数据层层解耦合,支持隐匿查询、隐私求交、联合统计、联合建模等多种服务扩展能力;支持从两方到任意数据参与方的动态加入;支持多种部署方式,计算引擎与数据接入可以一起部署也可以分开部署,系统管理可以部署在一方也可以部署在多方。灵活的部署方式和丰富的场景能力可以广泛满足多个参与方对联合统计、联合查询、AI建模等业务场景的需求。



油气管网数据安全共享平台参与方框架如图2所示。服务运营方引入服务开放方、服务审查方等，构建数据服务、算力引擎和运营管理等功能模块，在保证安全

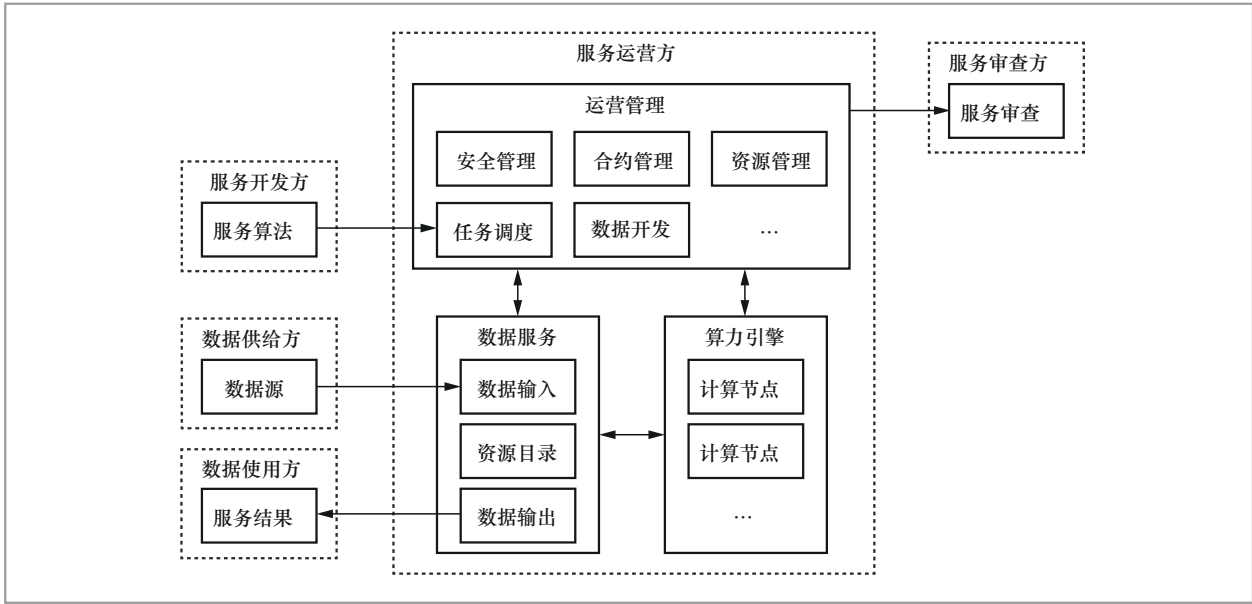


图2 油气管网数据安全共享平台参与方框架

⑧ 服务结算：经各方确认后，服务运营方进行服务结算。

⑨ 审查备案：服务结束后，服务审查方根据收集的服务合约、日志存证等对服务过程进行审查，并将审查结果和相关资料进行备案。

2.2 算法原理

油气管网数据安全共享平台根据服务的性质可以适配多种算法，而安全求交、隐匿查询、联邦学习及联合统计4种算法在实际应用场景中使用最广泛。后文对4种算法展开详细介绍。

2.2.1 安全求交

安全求交（private set intersection, PSI）是基于密码学的安全技术之一，在不泄露数据提供方和数据使用方用户名单的前提下协同计算并输出交集用户，但是各自无法获知交集以外的对方集合的任何信息。

目前业界有几种常用的PSI^[19]实现技术。其中，基于迪菲-赫尔曼（Diffie-Hellman, DH）密钥交换的PSI方案适用于小数据集的场景；基于无意识传输（oblivious transfer, OT）的PSI方案^[20-21]适用于大数据集的场景，但其系统的通信开销较大；基于不经意伪随机函数

（oblivious pseudorandom functions, OPRF）^[22]的PSI方案，在计算开销和通信开销方面均具有良好的表现。

基于OPRF的PSI方案如图4所示。客户端和服务端提前协商好一个伪随机数生成函数 $F(x, s)$ ，其中 x 是用于生成伪随机数的原始数据， s 是种子。客户端拥有一条原始数据 m ，当客户端和服务端执行OPRF协议后，服务器生成了一个种子 s ，客户端在未知种子 s 的情况下得到了伪随机数结果 $F(x, s)$ ，并且服务器未得到关于客户端的数据 m 的任何信息。类似OT协议，OPRF协议也是利用公钥密码体系来构造的，批量的OPRF协议可由批量OT实现的OT扩展协议来构造。类似OT扩展协议利用少量OT实现大量OT而减轻计算开销，批量的OPRF协议中，客户端可以利用少量的公钥复杂度计算和大量开销很小的对称密钥计算，得到大量的OPRF结果，其安全性仍然基于OT协议和伪随机数生成函数的对称密钥级别的安全性。

PSI方案利用布谷鸟哈希算法和批量OPRF协议，将客户端和服务器的所有数据映射为若干伪随机函数（pseudorandom function, PRF）结果，最终由客户端对比服务器发送过来的PRF结果确定哪些元素在两方的交集中，保护两方交集外的数据不泄漏。

2.2.2 隐匿查询

隐匿查询（private information retrieval, PIR）即隐私信息检索，查询方隐藏查询的关键词，数据持有方提供对应的加密查询结果，同时不暴露额外的数据信息，广泛应用于数据库查询、位置服务等领域，为用户提供了一种安全、私密的信息检索方式。

基于OPRF的PIR算法是最常用的隐



图4 基于OPRF的PSI方案

匿查询算法^[22]，其复用两次基于 OPRF 协议的 PSI 方案，实现隐匿的关键词查询，即客户端利用关键词向服务器数据对 (keyword-data) 进行查询。若客户端待查询的关键词处于服务器的数据库中，则查询得到相应的数据，而服务器对客户端待查询的关键词的任何信息未知，包括此关键词是否在服务器的数据库中。对于批量查询，基于 ORRF 的 PIR 方案具有安全高效的优势。

基于 OPRF 的 PIR 算法是当前最为广泛应用的隐匿查询算法之一^[22]。该算法巧妙地复用两次基于 OPRF 协议的 PSI 方案，以实现高效的隐匿关键词查询机制。具体而言，客户端利用特定的关键词向服务器发起查询请求，该服务器存储有关键词-数据对。若客户端查询的关键词确实存在于服务器的数据库中，则服务器将返回与该关键词相关联的数据信息。在整个查询过程中，服务器无法获取关于客户端所查询关键词的任何信息，同样无法获悉该关键词是否确实存在于其数据库中。对于需要执行批量查询的应用场景，基于 OPRF 的 PIR 方案展现出了卓越的安全性和高效性优势。

客户端与服务器之间需执行两次 PSI 协议，首次 PSI 协议的执行旨在使客户端能够确认其待查询的关键词是否存在于服务器的数据库中，并确定该关键词对应的 PRF 结果在服务器传输的 PRF 结果集合中的具体位置。此过程严格遵循 PSI 协议的安全性要求，确保服务器无法从中推断出客户端数据的任何信息。随后进行的第二次 PSI 协议，服务器会将数据与 PRF 结果进行混淆处理。客户端仅能依据首次 PSI 协议中获取的 PRF 结果位置信息，利用这些位置上的 PRF 结果与服务器进行比对。得益于 OPRF 协议的安全性保证，客户端

无法获取服务器的种子，仅能获得其自有数据的固定 PRF 结果，而无法通过计算获取服务器 PRF 结果的任何信息，因此，客户端无法推导出其他位置的 PRF 结果。对于其他位置的 PRF 结果所混淆的数据，在客户端看来均呈现为随机数据形态，从而有效保护了原始数据不被泄露。

2.2.3 联邦学习

联邦学习 (federated learning, FL)^[23] 又称联邦机器学习，是机器学习的一个框架，旨在解决分布式数据集的模型训练和模型推理的问题，是一个分布式训练方法。2016 年 McMahan 等^[24]首次提出了联邦学习这一术语，这项技术的主要思想是，在中心服务器的协调下各设备端之间相互协作训练模型，每个设备端的原始数据流都存储在本地，而无须进行数据的交换或传输，因此联邦学习具有隐私保护的属性，在设备端参与训练时，初始化模型会被发送到设备端，由于各设备端设备数据的差异，训练的模型参数权重会有不同，将各模型参数权重发送至中心服务器，中心服务器接收参数权重后，执行参数权重聚合操作来完成参数的更新任务，反复迭代上述操作直至模型收敛。在整个联邦学习的过程中，各参与方的数据始终保存在其本地服务器，降低了数据泄露的风险。

图 5 为联邦学习架构。联邦学习根据不同参与方的数据特点，可分为 3 类：横向联邦学习、纵向联邦学习和联邦迁移学习。

2.2.4 联合统计

联合统计是一种数据协作方法，多个数据持有方在确保不泄露各自原始数据隐私的条件下，共同执行统计分析 with 计算任务。数据使用者仅能获取到运算后的结果，

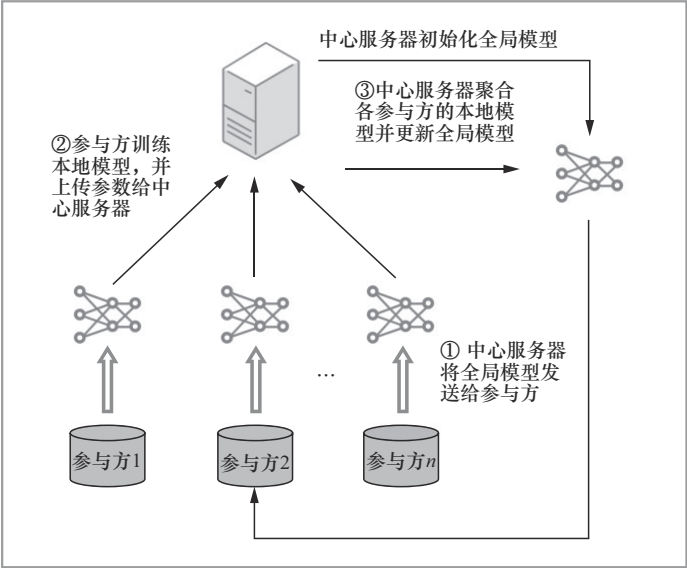


图5 典型联邦学习架构

而无法访问任何其他附加信息，算法的参数对平台本身及所有其他参与方均保持机密状态。通过这种联合处理方式获得的运算结果与直接在未加密的明文数据上进行运算所得的结果是一致的。

基于多方安全计算的联合统计方法可分为如下4个步骤：首先，采用多方计算协议实现基础运算操作；然后，利用基础运算操作组合实现衍生运算操作；之后，利用基础运算操作和衍生运算操作实现常用的函数和算法；最后，通过常用基础运

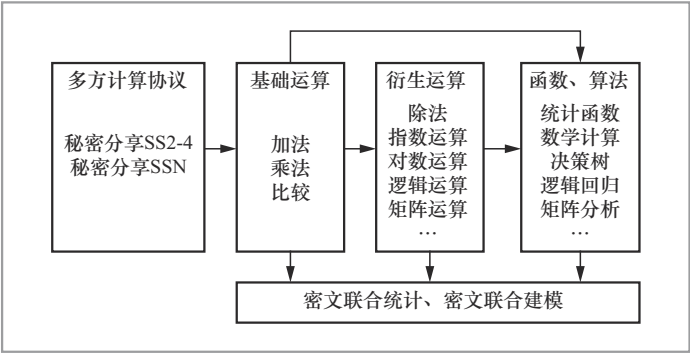


图6 联合统计计算流程

算、衍生运算和函数和算法组合实现联合统计计算。其各步骤间的关系如图6所示。

其中，多方安全计算协议采用秘密分享算法，其核心在于将一个秘密 S 拆分成 n 份，并分发给 n 个不同的参与方管理，只有将所有参与方结合才能恢复出原始的秘密 S ，而任意一个或多个参与方无法得知秘密的任何信息，在保证数据安全性的同时完成了数据的协同运算，实现数据的安全共享。

3 油气管网数据安全共享应用

不同数据安全共享技术的性能优劣各有不同，油气管网行业结合数据流通体量、安全性要求、使用场景等因素，灵活组合各技术，实现技术资源优势互补。

油气管网属于国家关键信息计算设施，产生的数据涉及社会民生，具有高度敏感性，在内部数据交换、外部数据共享以及与其他行业数据进行交互应用时，需对数据的安全性做到最大限度的保障。通过分析形成油气管网行业内部数据共享、外部数据利用、数据对外开放3个业务场景，本文梳理挖掘出可能结合的典型的数据安全共享应用场景，见表1。

本文选取油气管网行业3个典型的业务场景，建立油气管网数据安全共享平台，结合联邦学习、安全求交等技术拓展应用。

3.1 冬季保供调度优化

天然气运输过程会经过多个区域和省份，并且可能会出现突发的大量供气需求，需要在最短的时间内从供气方获取天然气。随着“全国一张网”的建设和运营，迫切需要油气管网运营单位统一协调各个区域

表1 油气管网数据安全共享计算场景

类别	分类依据	场景名称	子场景	场景描述	隐私计算算法				预期效果
					安全 求交	隐匿 查询	联邦 学习	联合 统计	
内部数据共享	企业内部各省网公司调度运输能力等数据敏感，跨地域分享困难	智能运维	管道事故模型构建	联合各公司的管道故障数据，训练故障分类模型等，分析疑难故障的种类和原因			☑		相比于单个企业，模型的准确度和精确度更高
			管道失效率预测	联合各公司的管道维修数据，训练管道失效率模型等，对未来特定时间周期管道失效率进行预测			☑		
		“全国一张网”	省网运营优化	对地区及省网公司的运营数据进行联合分析挖掘，提升运营指标	☑			☑	企业获取更准确的统计数据
外部数据利用	企业外部信息联动辅助决策，如运营商信息、地理气象信息、个人征信等涉密数据，难以公开	智能调度	多能综合调度	结合电力/风力等能源的调度数据，错峰调度			☑		实现动态资源规划和调度规划，保障供气
			冬季保供/应急优化	结合民用、商业和工业需求数据，配送公司运输能力，上游储备总量等，求解最优调度方案				☑	
		管道安全运维	管道线路设计规划	管道选线包括征地、高后果区、穿越区、气源、需求方等众多因素，涉及多个企业间的数据购买、采集、处理、传输和应用等，需对敏感数据尤其是地理数据进行保护		☑			保护油气管网数据不泄漏给第三方
			管道周边异常信号监测	结合电信运营商数据，对管道周边到访人数进行实时统计，并对疑似打孔盗油等异常情况进行预警		☑			为防范打孔盗油提供新的解决思路
		市场综合评级	客户风险评估	结合企业金融数据、政务数据等，对现有客户从准入到退出作出动态风险评估监测；同时也可以对缺乏历史信息的新客户，进行评分定级，识别潜力高价值客户	☑		☑		将手工定级（核验）升级为自动评级（核验），并且提高定级评估准确度
			供应商资质评价	结合外部数据如公安、税务、征信机构等，验证供应商资质和风险	☑		☑		
			供应链金融	打通供应链上下游企业数据，核对融资企业合同信息并评估企业融资风险		☑	☑		
数据对外开放	企业数据对外开放，形成统一的数据产品，发挥数据要素的乘数效应，由于同行竞争等原因，企业一般不公开	数据运营	数据安全共享交易平台	形成油气管道行业的数据资产和数据服务产品，用管网数据、标签、模型等数据资产为上下游企业赋能	☑	☑	☑	☑	打造数据资产，共享数据价值

和省份的天然气调度工作来保障冬季保供等场景。

冬季保供的关键环节是站点、管道线路的运输调度。一个典型的多方数据共享场景是，天然气运输经过多个区域，某一区域可能会突然出现大量供气需求，需在最短的时间内从供气方获取天然气。而区域管道企业、省网公司以及下游城市燃气配送公司都有各自的信息系统和数据资源库，站点和管道的工程建设、生产运维数据分管在各地，调度分析过程需要油气管道运营企业结合各自的天然气管道线路运输能力、供应商的供给量、托运商的需求量等数据，如区域管道企业及省网公司的天然气储备、城市燃气配送公司的运输能力，以及民用、商业、工业的天然气需求量等数据，共同进行分析统筹以便提供更好的优化调度方案。但跨区域、跨省天然气供销数据涉及调度运营和价值分配等因素，比较敏感，共享难度较大。

本文针对这一多方数据安全共享场景进行模拟和验证。A方作为天然气的供给方，为调度图的起点；下游城市燃气企业（D方）为需求方/应用方，为调度图的终点；其他方（B方、C方）为省网公司或配送公司，提供所在区域管道的运输能力，是调度的中间节点（站点）。几方联合运输站点和运输能力等数据，求解运输调度问题，得到最优调度线路。首先将上述问题简化为最短时间内输送的最大流问题。最大流是一类用于解决网络流问题的算法。网络流问题的核心是在一个有向图中，找到从源点（天然气存储方）到汇点（城市燃气用气方）的最大流量。图中的每条边都有一个容量，表示可以通过该边的最大流量。最大流算法的目标是确定各边上的流量分配，使从源点到汇点的总流量最大

化，实验方案的输入数据包含供给方站点（源点）的天然气供给量、需求方站点（汇点）对天然气的需求量、从起点到终点不同管道线路的走向，以及管道线路在方向上的最大运输能力。这些管道线路属于不同的区域管道企业和省网公司。输出结果为从源点到汇点的最大流量线路。

多方最大流调度流程如图7所示，关键业务流程如下。

①数据准备：D方作为调度发起方创建项目，A、B、C各方选择相应的数据。

②算法选择：为了简化实验，采用的算法是最大流算法（实际可以扩展到多种应用算法），它的优点是思路直观、性能优越、通用性强。

③联合计算：数据应用验证通过后，采用最大流算法在密文数据上进行交互和求解方程计算。

④结果生成：计算完成后生成最优调度路径。

在部分区域网络运输能力未知的情况下，可借助多方安全计算等技术，对运输调度问题进行求解，提供最优方案，从而实现综合智能调度优化方案，达到冬季保供和内部提质增效的目的，同时保护各方原始数据隐私不泄漏，打消机构间不愿共享数据的安全顾虑。

图8为冬季保供调度方案。通过油气管网数据安全共享平台接入部委等多个需求方的天然气需求量、供给方的天然气储备总量、省网公司的线路运输能力等外部数据，依托联合统计算法在密文下求解最优调度方案，统一运营企业解密结果，得到最终的调度规划线路。

3.2 管道安全运维提升

油气管网运营企业维护的管道存在距

坐标等敏感信息，也不能对外泄漏。本文借助多方安全计算技术保护各方数据隐私，通过隐匿查询算法，对管道地理信息与运营商运营地理数据进行叠加运算，油气管网运营企业隐藏具体的地理位置信息，获取该位置附近的实时到访人数统计数据，并与历史正常人口活跃度进行对比，如果超过正常阈值进入监视预警。隐匿查询算法实现流程如图9所示，油气管网运营企业为到访人数查询方，数据提供方能够提供基础地理数据、区域人流数据。

在管道周边防盗油信号检测场景中，有 N 个运营商，每个运营商可提供空间区域内的到访人口统计数据。油气管网运营企业将管道线路作为查询条件，在向运营商隐藏管道线路的地理位置前提下，运营商对管道线路图层和自身的人流图层在加密状态下做空间叠加计算，实时统计管道周边的到访人口数量，并按业务要求提供分钟级颗粒度汇总数据，油气管网运营企业得到的结果仅为管道周边的人口访问数量（活跃度）。主要实现过程如下。

①数据准备：查询方作为调度发起方创建项目/应用，被查询方（数据提供方）选择相应的数据。

②算法选择：由于需要隐匿查询方的查询条件和意图，各方选择隐匿查询PIR算法。

③隐匿查询：查询方输入查询条件，该条件加密后与被查询方的实时数据进行匹配。此处可设置自动查询方式和查询间隔时间。

④结果生成：查询方获得匹配结果后，可对选定时间（如10 min）内出现的流动人口活跃信号超过阈值进行预警。通过在地图上绘制历史查询数据的热力图，得到一天内管道周边信号的波动曲线。查询过程中通过与历史趋势曲线进行对比，可判断是否出现异常情况，并对异常情况进行监测，从而实现对安全运维数据的交叉验证。

如图10所示，油气管网运营企业通过油气管网数据安全共享平台接入电信运营商的区域人流数据，通过输入查询条件，如某段管道的GIS或者网格数据，依托隐

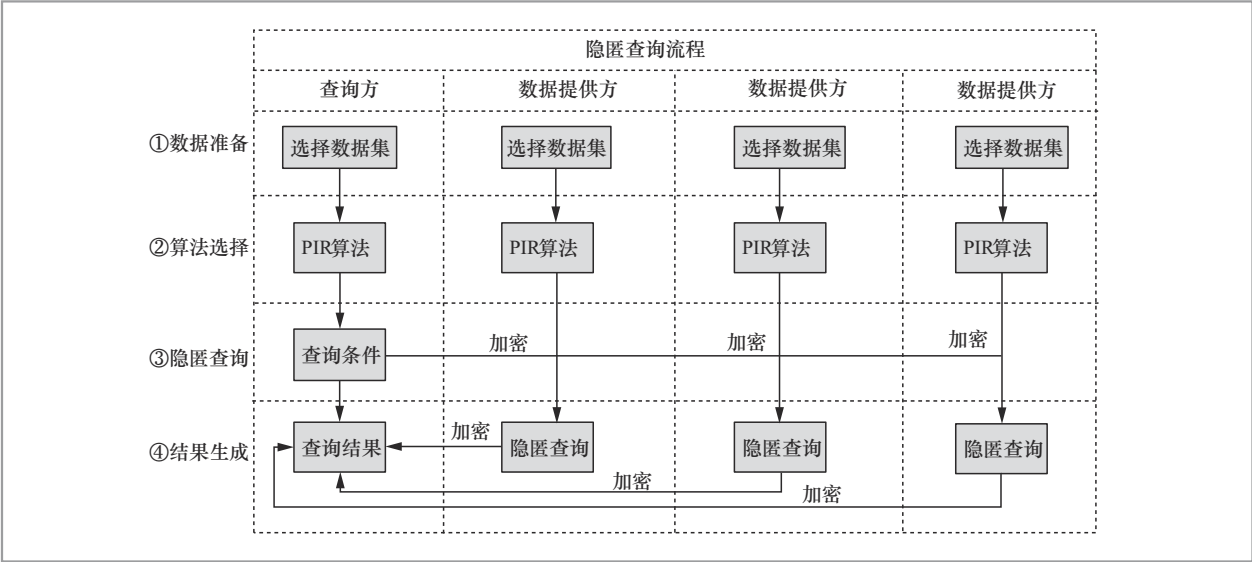


图9 隐匿查询算法实现流程

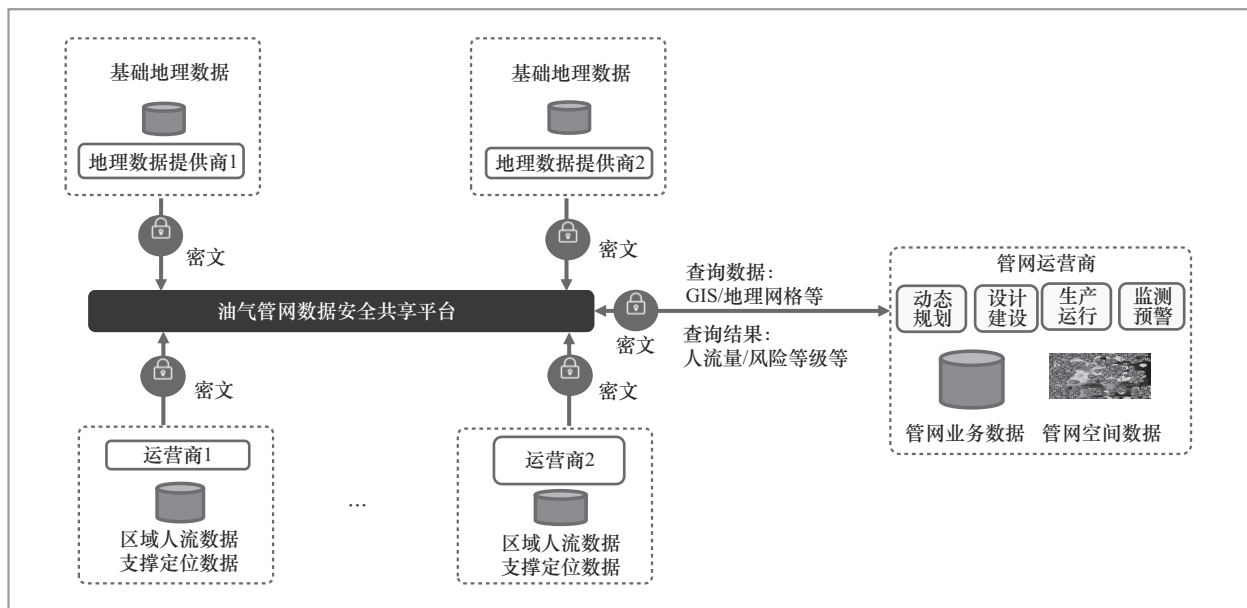


图10 管道安全运维监测

匿查询算法从电信运营商的数据库中查得实时的人流活跃度数量。油气管网企业解密结果，并将查询结果与历史正常人口活跃度进行对比，如果超过正常阈值就进入监视预警。如果这种情况持续一段时间，大致可判断有异常情况，采取下一步防范措施。

基于隐私计算的油气管网数据安全共享技术给防范打孔盗油提供了一种新的解决思路，创造性地结合大数据、隐私计算等前沿技术对打孔盗油进行远程分析检测以及实时预警，加强防范手段，节省人力投入成本以及监控设备的采购费用；同时，帮助运营商减少对数据泄漏的顾虑，从而推动运营商开放更多更准确的数据，将其更好地应用到管道的安全生产和运维中。

3.3 市场客户综合评级

当前企业客户工作缺乏系统化管理，只有点的覆盖，没有构筑全面的、立体的

客户关系。客户从准入到退出，企业缺乏主动挖掘客户需求的意识，没有分析、排序、分发、跟踪、确认关闭的全过程管理。对于新客户，由于缺乏历史数据，企业无法对客户进行评级，对于出现经营异常的客户预警能力不足。

油气管网运营企业在已有客户历史签约金额、合同执行等数据的基础上，对客户的管理引入银行、征信机构、公安部门、政府等多方信息，在不收集、不存储原始数据的前提下，保障数据不出本地、安全合规共享的基础上，使用联邦学习技术进行模型训练，并结合联合统计方法制定业务规则，将训练好的模型对新客户进行动态预测，获得准确的客户评分和客户全景风险信息，既在KYC、客户特征勾勒、黑名单信息共享上发挥重要作用，又保护了油气管网运营企业的客户信息不泄漏，如图11所示。

在市场综合定级的场景中，有 N 个参与方，每个参与方提供用户不同的特征数

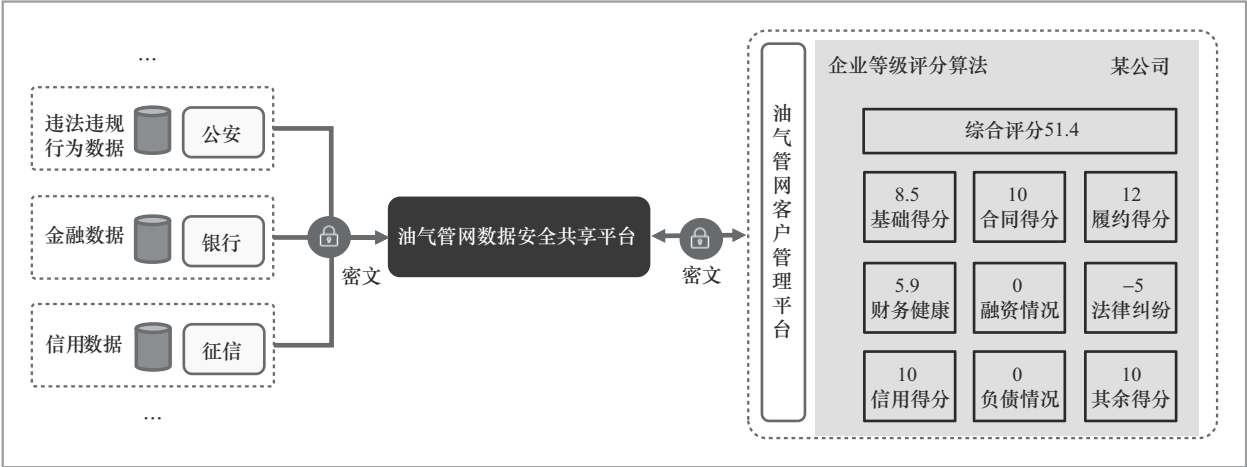


图11 市场客户定级评估

据。如油气管网运营企业客户合同数据、银行金融数据、征信机构征信数据、公安司法数据等。**图12**是使用联邦学习技术实现油气管网运营企业和银行两方数据的安全共享，在不直接共享原始数据的情况下，协同训练出市场客户综合评级模型。

如**图12**所示，输入数据为油气管网企业客户历史合同特征和评分，以及该客户银行金融特征，两方联合客户的不同特征维度训练出客户综合评级模型，输出结果为客户的信用评分值，评分越高代表风险（违约率）越低。企业市场客户综合评级模

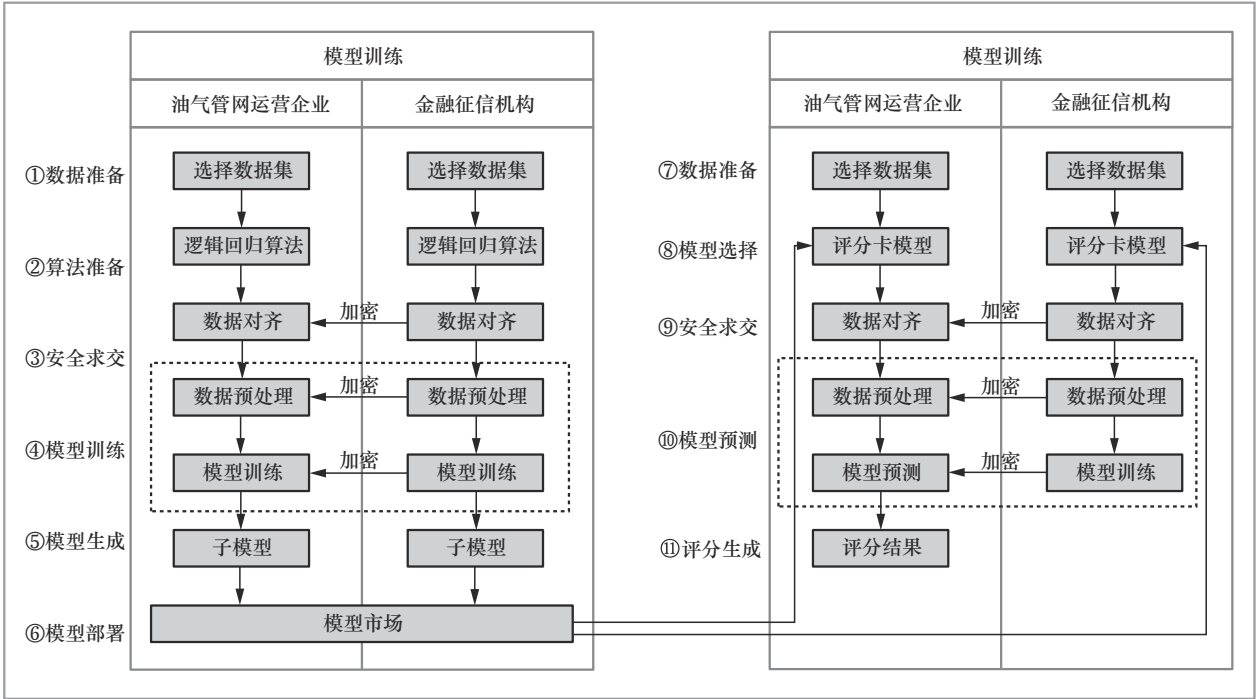


图12 两方评分卡联合建模数据流程

型训练的关键流程如下。

①数据准备：应用方作为模型发起方创建项目，选择数据集，油气管网运营企业提供特征和标签，金融征信机构提供特征。

②算法准备：构建企业评分卡模型使用逻辑回归算法，它的优点是原理简单、调参简单，且模型结果的可解释性强，可以更好地转化为量化的分数。

③安全求交：两方的数据应用验证通过后，采用 PSI 算法进行安全求交。安全求交允许持有各自数据集的双方在不直接交换或泄露各自原始数据的前提下，共同计算出交集用户。

④模型训练：在对齐的数据上设置数据预处理规则，包括数据分箱和 WOE 编码等。数据预处理完成后，开始模型训练，设置逻辑回归模型训练参数包括迭代轮数、学习率等。模型训练完成后，进行模型评估，用户可以根据评估指标（如 AUC、KS、Loss、Accuracy 等）判断模型是否合理，对不合理的模型调整数据配置和算法后，重新训练。

⑤模型生成：模型训练完后，每方得到相应的子模型。

⑥模型部署：已训练完成的模型可以再导出或发布至模型市场。

客户评分预测的关键步骤如下。

⑦数据准备：油气管网运营企业在预测客户市场定级时，选择预测数据集（不含标签）。

⑧模型选择：选择模型训练生成的模型。

⑨安全求交：两方采用 PSI 算法进行安全求交。

⑩模型预测：数据对齐后可以设置数据预处理规则，包括数据分箱和 WOE 编码，两方数据在模型下进行计算；

⑪结果生成：计算完成后，生成预测结果评分，评分越高代表客户违约率越低，客户风险越低。

以上方法将联邦学习技术应用到管网市场客户综合定级中，既保护了油气管网运营企业的客户信息不泄漏，又为客户全生命周期管理提供动态风险监测。对新用户和现有用户进行动态定级，识别潜力高价值客户，更好地服务客户，提高客户满意度，进一步丰富客户画像，解决对客户认识不全面、经营异常情况预警能力不足的问题。

4 结束语

本文基于隐私计算的油气管网数据安全共享技术为油气管网运营企业数据的共享和利用提供技术支持，致力于推进油气储运行业协同发展，推动油气管网数据资源的流通和使用，打通不同主体间的壁垒，带动油气管网一张网各环节的互联互动互补，提升产业链上下游及行业间协调运行效率，打通跨域数据的应用价值链，使数据在各业务线之间安全共享和流通，为高价值、高成本、高敏感的行业数据释放数据要素价值提供了可能性。

未来，笔者还将继续基于多方安全计算、联邦学习、可信计算、区块链等技术，一站式完成油气管网行业数据确权、分类分级、数据定价、数据监管、交易撮合、数据市场、交易大厅、数据资产凭证等创新型数据管理与流通方式，通过部署行业数据安全共享平台，实现数据流通的支撑性技术，构建能源行业数据流通的基础性设施，进一步结合发挥智能管网的优势，推动形成智能调控体系，实现资源的精准高效配置。

参考文献：

[1] 黄维和, 宫敬. 天然气管道与管网多能融合技术展望[J]. 油气储运, 2023, 42(12): 1321-1328.

HUANG W H, GONG J. Prospect for the development of natural gas network and the multi-energy integration technology in pipeline networks[J]. Oil & Gas Storage and Transportation, 2023, 42(12): 1321-1328.

[2] 高振宇, 张慧宇, 高鹏. 2022年中国油气管道建设新进展[J]. 国际石油经济, 2023, 31(3): 16-23.

GAO Z Y, ZHANG H Y, GAO P. New progress in China's oil and gas pipeline construction in 2022[J]. International Petroleum Economics, 2023, 31(3): 16-23.

[3] 乔士航, 朱梦茹. 中国油气智慧管网的现状与发展[J]. 化工管理, 2024(5): 18-21.

QIAO S H, ZHU M R. Current situation and development of China's oil and gas smart pipeline network[J]. Chemical Enterprise Management, 2024(5): 18-21.

[4] 吕晓波, 姜刚. 基于物联网数据分析的智能化油气管道系统建设[J]. 石油石化物资采购, 2024(4): 206-208.

LYU X B, JIANG G. Construction of intelligent oil and gas pipeline system based on data analysis of Internet of Things[J]. Petroleum & Petrochemical Material Procurement, 2024(4): 206-208.

[5] 姜昌亮. 油气管道全生命周期质量管控与安全管理探讨[J]. 油气储运, 2023, 42(10): 1081-1091.

JIANG C L. Quality control and safety management of oil and gas pipelines during its whole life cycle[J]. Oil & Gas Storage and Transportation, 2023, 42(10): 1081-1091.

[6] 黄维和, 韩景宽, 王玉生, 等. 我国能源安全战略与对策探讨[J]. 中国工程科学, 2021, 23(1): 112-117.

HUANG W H, HAN J K, WANG Y S, et al. Strategies and countermeasures for ensuring energy security in China[J]. Strategic Study of CAE, 2021, 23(1): 112-117.

[7] CHENG W, CHENG Y F. A techno-economic study of the strategy for hydrogen transport by pipelines in Canada[J]. Journal of Pipeline Science and Engineering, 2023, 3(3): 100112.

[8] 黄维和, 张曦, 张元涛, 等. 天然气下游大数据技术研究与应用综述[J]. 油气与新能源, 2021, 33(4): 1-5.

HUANG W H, ZHANG X, ZHANG Y T, et al. Review on research and application of big data technology in natural gas downstream market[J]. Petroleum and New Energy, 2021, 33(4): 1-5.

[9] 安小米, 郭明军, 洪学海, 等. 政府大数据治理体系的框架及其实现的有效路径[J]. 大数据, 2019, 5(3): 3-12.

AN X M, GUO M J, HONG X H, et al. Framework of government big data governance system and effective way of implementation[J]. Big Data Research, 2019, 5(3): 3-12.

[10] 于宏源, 关成龙. 美国克罗尼尔安全事故对我国能源安全韧性的警示[J]. 能源, 2021(6): 26-29.

YU H Y, GUAN C L. Warning of Cronier safety accident to China's energy security and resilience[J]. Energy, 2021(6): 26-29.

[11] 冯庆善. 智能油气管网系统建设与运行方法论研究[J]. 油气储运, 2024, 43(8): 841-854.

FENG Q S. Research on construction and operation methodology for intelligent oil and gas pipeline network systems[J]. Oil & Gas Storage and Transportation, 2024, 43(8): 841-854.

- [12] 吴炳泽. 油气智慧管网运维管理系统设计与应用研究[D]. 重庆: 重庆大学, 2022.
- WU B Z. Research on design and application of operation and maintenance management system for oil and gas smart pipe network[D]. Chongqing: Chongqing University, 2022.
- [13] 方滨兴, 贾焰, 李爱平, 等. 大数据隐私保护技术综述[J]. 大数据, 2016, 2(1): 1–18.
- FANG B X, JIA Y, LI A P, et al. Privacy preservation in big data: a survey[J]. Big Data Research, 2016, 2(1): 1–18.
- [14] 余晗, 梁音, 宋继勳, 等. 数据安全共享技术发展综述及在能源电力领域应用研究[J]. 信息安全研究, 2023, 9(3): 208–219.
- YU H, LIANG Y, SONG J M, et al. A survey of data security sharing technology development and its application in power domain[J]. Journal of Information Security Research, 2023, 9(3): 208–219.
- [15] 王健宗, 孔令炜, 黄章成, 等. 联邦学习隐私保护研究进展[J]. 大数据, 2021, 7(3): 130–149.
- WANG J Z, KONG L W, HUANG Z C, et al. Research advances on privacy protection of federated learning[J]. Big Data Research, 2021, 7(3): 130–149.
- [16] 陈长松. 零信任架构下的数据安全纵深防御体系研究[J]. 信息网络安全, 2021, 21(S1): 105–108.
- CHEN C S. Research on defense-in-depth system of data security under zero trust architecture[J]. Netinfo Security, 2021, 21(S1): 105–108.
- [17] 祝烈煌, 董慧, 沈蒙. 区块链交易数据隐私保护机制[J]. 大数据, 2018, 4(1): 46–56.
- ZHU L H, DONG H, SHEN M. Privacy protection mechanism for blockchain transaction data[J]. Big Data Research, 2018, 4(1): 46–56.
- [18] 龚俭, 臧小东, 苏琪, 等. 网络安全态势感知综述[J]. 软件学报, 2017, 28(4): 1010–1026.
- GONG J, ZANG X D, SU Q, et al. Survey of network security situation awareness[J]. Journal of Software, 2017, 28(4): 1010–1026.
- [19] MEADOWS C. A more efficient cryptographic matchmaking protocol for use in the absence of a continuously available third party[C]//Proceedings of the 1986 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 1986: 134.
- [20] KOLESNIKOV V, KUMARESAN R, ROSULEK M, et al. Efficient batched oblivious PRF with applications to private set intersection[C]//Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2016: 818–829.
- [21] BENNY P, THOMAS S, MICHAEL Z. Faster private set intersection based on OT extension[C]//Proceedings of the 23rd USENIX Security Symposium. Berkeley: USENIX Association, 2014: 797–812.
- [22] BONEH D, KOGAN D, WOO K. Oblivious pseudorandom functions from isogenies[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020: 520–550.
- [23] MCMAHAN H B, MOORE E, RAMAGE D, et al. Federated learning of deep networks using model averaging[EB]. arXiv preprint, 2016, arXiv: 1602.05629.
- [24] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[EB]. arXiv preprint, 2016, arXiv: 1602.05629v4.

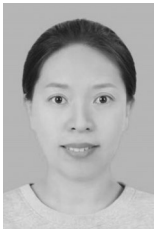
作者简介



任武（1984—），男，国家石油天然气管网集团有限公司科学技术研究总院分公司高级工程师，主要研究方向为油气管网数据资产要素化和数据安全等。



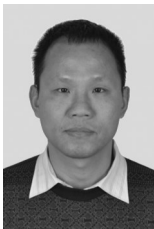
张珂（1998—），女，国家石油天然气管网集团有限公司科学技术研究总院分公司工程师，主要研究方向为油气管网数据和网络信息安全、隐私保护等。



富影杰（1981—），女，国家石油天然气管网集团有限公司东北分公司沈阳油气计量中心工程师，主要研究方向为油气计量、数据管理。



张正雄（1995—），男，国家石油天然气管网集团有限公司科学技术研究总院分公司工程师，主要研究方向为数字化、北斗时空数据管理。



蔡永军（1980—），男，国家石油天然气管网集团有限公司科学技术研究总院分公司高级工程师，主要研究方向为油气管网安全状态监测、数字化转型、智能化升级。

收稿日期: 2024-08-02
通信作者: 任武, renwu@pipechina.com.cn
基金项目: 国家管网项目 (No.ZYZL-SZZN-202212)
Foundation Items: National Pipeline Project (No.ZYZL-SZZN-202212)