科技评论

# 数据要素时代的隐私计算平台互联路径与挑战

刘颖慧[1,2],魏进武[2*],张溶芳[2],蔡一欣[2],李堃[2]

**摘要** 隐私计算是实现“数据可用不可见”的核心技术,隐私计算平台繁荣发展,逐渐形成了以隐私计算平台为核心的“平台孤岛”现象。如何解决“平台孤岛”问题,进一步促进数据要素融合和价值成为重要的研究方向。梳理了国内外典型的隐私计算平台功能、国内隐私计算平台的互联生态现状,结合行业实践案例,明确节点互联和算法互联的研究范畴,剖析节点和算法互联的不同模式。总结了隐私计算跨平台互联的算法、数据、安全和性能难点,虽然隐私计算平台互联互通面临底层技术不兼容、跨平台数据资源操作难和数据泄露风险高,以及多方通信性能难以满足用户需求等多种挑战,但是跨平台互联仍然是未来发展的重要趋势。标准和规范方面,金融行业、电信行业都在探索行业内部的隐私计算平台互联标准。技术方面,双向适配模式下的双方适配功能将逐渐被“适配器”式的转换产品替代,而规模化的、基于开放协议的平台互联互通将逐渐成为更多技术方的选择。

**关键词** 数据要素;隐私计算;节点互联;算法互联

随着数据要素被确认为国家最新的生产要素,如何加速数据的要素化、市场化配置进程成为国家重点关注的问题。《关于构建数据基础制度更好发挥数据要素作用的意见》《数字中国建设整体布局规划》《“数据要素×”三年行动计划(2024—2026年)》等系列国家政策的发布,体现了国家对数据要素“供得出、流得动、用得好”最新工作要求。“供得出”强调数据的开放与共享,而“流得动”与“用得好”则是数据产生价值倍增效应的关键环节。保障数据安全与隐私保护是数据流通与应用的底线要求,因此,隐私计算成为在保护隐私的基础上进行多方数据协同的关键技术之一。

隐私计算按照当前行业的主流思想可以分为3类技术,安全多方计算(secure multi-party computation, MPC)、联邦学习(federated learning, FL)和可信执行环境(trusted execution environment, TEE)。安全多方计算和联邦学习是解决数据持有方不愿开放或受限开放数据的前提下,如何实现安全的多方协作计算、联合建模等需求的关键技术,而可信执行环境通过隔离的方式保障了数据和计算任务的安全。因此,隐私计算是实现“数据可用不可见”的热门且核心技术之一。随着隐私计算技术的蓬勃发展,通过数据加密、参数共享代替数据共享等模式,实现了“数据孤岛”的有效连接,数据持有方的数据在不出其管理域的前提下,完成了多方协作计算。但是,由于软件厂商的隐私计算平台技术框架、管理模式、算法实现和资源调度模式都不尽相同,“数据孤岛”虽然在一定程度上实现了连接,新的、以软件为核心的“平台孤岛”还是产生了。各个数据持有方只有采用相同的计算平台和框架才能进行多方计算任务,在部署

1. 北京交通大学电子信息工程学院,北京 100044
2. 中国联合网络通信有限公司研究院,北京 100176

收稿日期:2024-08-21;修回日期:2025-06-06
作者简介:刘颖慧,高级工程师,研究方向为数据流通,电子信箱:liuyh17204@chinaunicom.cn;魏进武(通信作者),教授级高级工程师,研究方向为卫星互联网与大数据,电子信箱:weijw@chinaunicom.cn
引用格式:刘颖慧,魏进武,张溶芳,等. 数据要素时代的隐私计算平台互联路径与挑战[J]. 科技导报, 2025, 43(24): 17-26; doi:10.3981/j.issn.1000-7857.2024.08.01007

多个框架、重新适配业界主流框架和推动不同平台框架的互联开放这 3 个选项中,从部署成本、扩展性和便利性来看,推动不同平台框架的互联开放将成为解决“平台孤岛”问题的最优方案,如何实现不同隐私计算框架的互联成为了业界关注和亟待解决的难题。

## 1 隐私计算平台互联现状

### 1.1 隐私计算平台发展现状

根据公开信息,整理了国内外主流的隐私计算平台和典型隐私计算工具,隐私计算平台从上文提到的主流技术视角主要分为 3 类,具体如表 1[1−15]所示。

一是以工具包、引用库形式提供隐私计算的必要功能。行业上相对成熟的主要包括数据匿名化、差分隐私、同态加密库等。典型应用包括德国柏林慈善大学主导研发的 ARX 匿名化开源工具包、国际商业机器公司(IBM)的 HElib 同态加密库、PALISADE 格密码开源库和谷歌公司推出的 TensorFlow Privacy 等。其技术核心都在于构建安全多方计算和联邦学习的基础能力和协议,并且采用与计算引擎松耦合的模式,实现跨平台的应用。与之对应的,则是与已有计算平台紧耦合的升级库,如 TF−Encrypted 和 Tensorflow Federated Framework(TFF),就是与谷歌公司的 Tensorflow 计算框架深度集成的升级隐私计算框架。

二是只提供多方安全计算、联邦学习或可信执行环境单一功能的平台。典型平台如美国伯克利大学 RISE 实验室(UC Berkeley RISE Lab)开发的安全多方计算框架 MC²、瑞士安全科技公司 De-

表 1 国内外隐私计算平台工具

| 平台/框架名称 | 独立工具包/库 | FL | MPC | TEE | 开发公司/机构 |
|---|---|---|---|---|---|
| ARX[1] | 匿名化 | | | | 德国柏林慈善大学 |
| Cape Privacy[2] | | √ | √ | | Cape Privacy |
| Decentriq | | | | √ | Decentriq |
| diffprivlib[3] | 差分隐私 | | | | IBM |
| ObliVM[4] | | | √ | | University of Maryland |
| Duet | 差分隐私 | | | | OpenMinded |
| Inpher XOR Platform[5] | | √ | √ | | Inpher |
| MC²[6] | | | √ | | UC Berkeley RISE Lab |
| SEAL[7] | 同态加密 | | | | 微软 |
| Private Join and Compute[8] | | | √ | | 谷歌 |
| PySyft[9] | | √ | √ | | OpenMined |
| Sharemind[10] | | | √ | | Cybernetica |
| TF−Encrypted[11] | 同态加密、安全多方计算协议 | | | | DropoutLabs, Openmined等 |
| Tensorflow Federated Framework[12] | | √ | | | Google |
| CrypTen[13] | | | √ | | Facebook |
| 点石隐私计算平台 | | √ | √ | √ | 百度 |
| FATE[14] | | √ | | | 微众银行 |
| SecretFlow隐语平台[15] | | √ | √ | √ | 蚂蚁集团 |
| 密流安全计算平台 | | √ | √ | | 翼支付 |
| 联通链隐私计算平台 | | √ | √ | | 联通数科 |
| 中国移动1+X隐私计算平台 | | √ | √ | | 中国移动 |
| InsightOne隐私计算平台 | | √ | √ | | 洞见科技 |
| 慧点隐私计算平台 | | √ | √ | | 招商银行 |
| 阿凡达平台 | | √ | √ | | 富数科技 |
| 基于隐私计算的共享智能平台 | | √ | √ | | 同盾科技 |
| PrivPy多方安全计算平台 | | | √ | | 华控清交 |

centriq 研发的基于可信执行环境的数据净室等应用。

三是提供多类功能的综合性平台。国内外多个平台提供的包含联邦学习、安全多方计算和可信执行环境中的多种能力。中国隐私计算的平台,以互联网公司、通信行业、金融行业和软件服务商为主要研发方,其多侧重于研发综合性隐私计算平台。在第二类和第三类平台中,平台研发方会研发自有匿名化、差分隐私和安全多方计算的各种基础安全协议,如同态加密、不经意传输、零知识证明等。当前,隐私计算的开源主要以互联网企业平台为主,如微众银行的 FATE(federated AI technology enabler)框架和蚂蚁集团的隐语(SecretFlow)平台都开源了其代码。

隐私计算平台与区块链技术的结合逐渐成为主流,在遵循安全协议的基础上,通过区块链进行身份认证、对计算参与方的行为数据进行上链和以智能合约实现协议自动化等方式,进一步保障计算参与者的身份可信和行为可追溯,如 InsightOne、阿凡达和联通数链隐私计算平台都采用了区块链技术进一步保障隐私计算参与方的身份与行为安全可信。

### 1.2 隐私计算平台互联的生态现状

隐私计算平台互联场景针对不同的数据持有方,采用不同的隐私计算平台进行联合建模或联合计算时,因计算框架或采用的计算工具不同进一步导致“平台孤岛”。根据调研结果,国外隐私平台互联互通的尝试较少,多是基于单个隐私计算平台的小范围内协同计算应用。中国的研究机构、互联网行业、金融行业都开始探索隐私计算框架的互联互通,总体生态如图 1 所示,从代码开源角度可分为开源互联和闭源互联。从互联的推动主体视角,主要包括以下 3 类。

一是以金融行业为中心的、代码闭源隐私计算平台互联互通。北京金融产业科技联盟是在中国人民银行指导下,由中国金融电子化集团发起的综合性金融科技联合创新工作平台,成员单位 380 余家。该联盟牵头开展了金融领域隐私计算数据流通平台的互联互通模式探索[16],成为金融行业的主流互联互通标准。该方案以管理面和数据面切分为设计思路,将隐私计算拆解为管理互通、流程调度互通、算法互通、算子互通、协议开放等关键技术。招商银行以其慧点隐私计算平台为核心,协同包含 InsightOne 平台、阿凡达平台在内的 4 家平台,构建了“可插拔平台框架+算法组件架构”的互联互通方案与实践[17]。

二是以互联网企业为中心的、开源研发生态。微众银行开源的 FATE 2.0 开放开源异构系统互联互通开发指南[18],意图推动异构系统的互联互通。蚂蚁集团开源的隐语框架,在隐私计算联盟的组织下,协同中国信息通信研究院(信通院)、电信运营商、银行业和服务商等,制定系列隐私计算跨平台

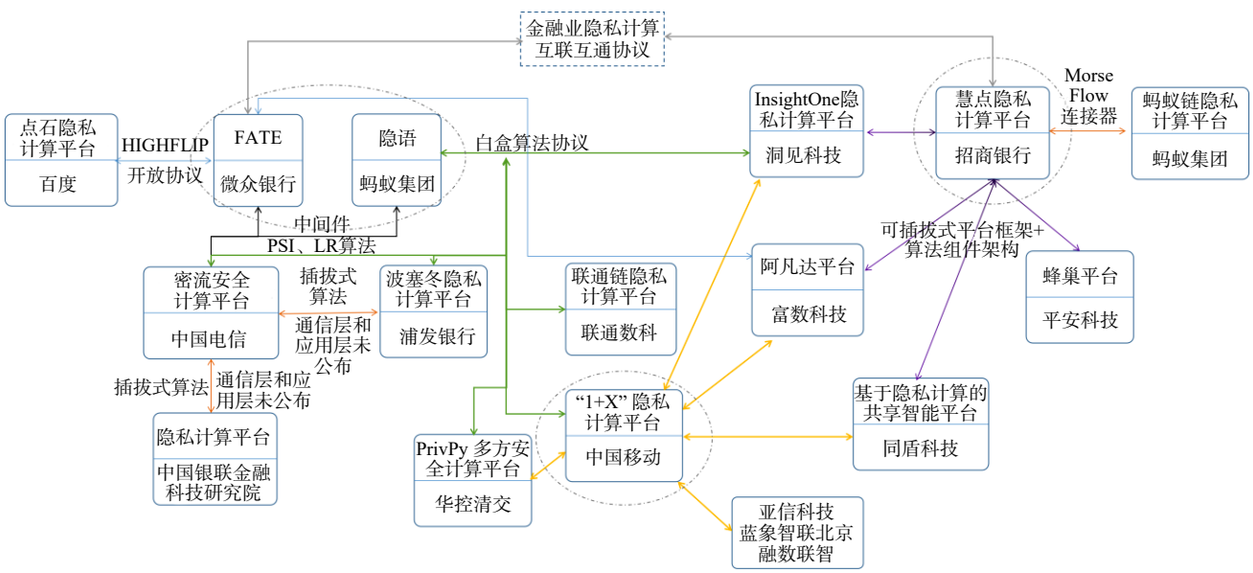


图 1 当前隐私计算平台互联开放生态

互联互通开放协议,并开放具体的传输层协议和基于椭圆曲线迪菲–赫尔曼密钥交换实现的隐私集合求交(elliptic curve diffie–hellman private set intersection, ECDH–PSI)、基于秘密分享实现的逻辑回归(secret sharing logistic regression federated logistic regression, SS–LR)[19]、基于半同态的联邦线性回归(partially homomorphic encryption, PHE–FLR)和安全梯度提升(secure gradient boosting, SGB)4 个算法的开放协议。

三是点状的以自研平台为中心推进的、闭源的互联互通实践。中国电信的密流安全计算平台,与浦发银行的波塞冬隐私计算平台和中国银联金融科技研究院的隐私计算平台构建“插拔式”算法组件的互联互通,与 FATE 和隐语构建基于“中间件”的隐私集合求交和逻辑回归算法互通。中国移动基于其“1+X”隐私计算平台与洞见科技、富数科技、同盾科技和华控清交等多个软件服务商公司实现互联互通[20]。

## 2 隐私计算平台互联的技术路径

不同隐私计算平台开发方就互联路径尚未达成共识。主要有 2 种思路,一是从平台架构和开放层次深度,在 2022 年可信隐私计算峰会上,有企业提出隐私计算互联从系统架构视角,可以按照应用层、算法层和原语层进行分层互通。应用层面向管理系统实现节点的发现、资源管理授权等互通,算法层实现算法、计算任务、模型、计算结果等系列协议互通,最底层实现通信、加密、随机、压缩、摘要等计算原语的协议互通。北京金融产业科技联盟提出按照管理面和数据面解耦进行互联互通[16],管理面指的是包含节点、数据、项目、流程、组件、作业所有进行计算任务需要的管理元素的互通,而数据面更多指的是实际的算法互通。叶剑等[21]提出按照“底层通信—中间层交互—顶层应用”的思路设计隐私计算平台的互联互通,通信层聚焦通信框架和机制的规范化,交互从节点、资源和算法 3 个维度进行全环境的流程状态互通,而应用层聚焦跨平台协同作业的任务调度、监控和存证。

### 2.1 研究范畴

中华人民共和国工业和信息化部批准发布的行业标准《隐私计算跨平台互联互通第 1 部分:总体框架》(编号 YD/T 4961.1—2024)定义隐私计算平台互联互通为“具有不同系统架构或功能实现方案的隐私计算技术平台(包括同一平台的不同版本)之间通过统一规范的接口、协议等实现跨平台数据、算法、算力的交互与协同,以支持部署不同技术平台产品用户共同完成同一隐私计算任务”。

隐私计算平台互联互通的目标是无需数据计算参与方重复部署不同隐私计算平台的前提下,部署轻量级的互通适配功能或修订隐私计算平台内部流程、机制和算法,实现不同的隐私计算平台可以完成跨平台的联合计算任务。因此,隐私计算平台互联互通应聚焦以下研究范畴和基准。

本文主要讨论在保障原平台的自治性基础上的互联。在推动隐私计算跨平台互联的过程中,不同框架的节点不能直接调度其他平台的所有计算资源,包括其数据集和最小粒度的计算单元。如具备调度全资源的能力,则破坏了本身系统自治性,本质上是构建了新的隐私计算统一平台。节点可以发现其他隐私计算平台的部分公开的代理计算节点,通过代理节点的互联互通间接完成计算单元的调度。算法互联也是同理,隐私计算平台只能调用自有平台的基础算法组件和算子,不能跨平台调度,但如果是外部第三方插件式的算法组件,多方都可调度算法相同算法组件实现互通。

需要指出的是,数据资源互联的不在本文讨论范围内。隐私计算的核心是“数据可用不可见”,即不会产生原始明文数据的跨域访问,在非平衡隐私求交等场景下,数据的传输也是采用了如全同态加密后的密文数据传输。不同隐私计算平台开放数据集以供互相访问的场景不在本文的探讨范围内。

本文隐私计算平台互联讨论内容如图 2 所示。当发生了需要跨隐私计算平台协同计算的需求,可以是一个项目,也可以是一个查询、统计需求等。产生协同计算需求后,按照需要数据的持有方,形成统一的作业(job)编排,通常以有向无环图(directed acyclic graph, DAG)表述,将计算作业拆解成不同的计算任务(task),并明确每个计算任务需要调用的组件和输入输出元数据。在计算过程

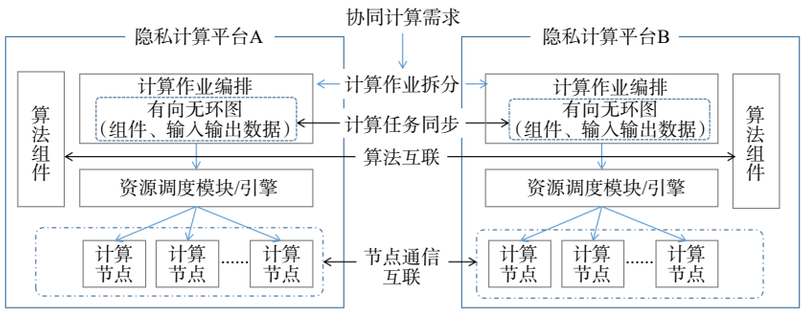


图 2 隐私计算平台互联研究内容

中,计算前的作业拆分结果信息和计算过程中的任务状态与结果同步都需要隐私计算平台的实际执行节点可以互相通信,节点互联在 2.2 节中详述。算法组件的互联是指实际计算过程中核心互联部分,在 2.3 节中详述。

## 2.2 节点互联

参与联合建模、联合计算等多方协同作业的参与方,记作 $P_i$,$i$ 表示参与方编号。节点是指承载具体计算任务的节点,可以是物理机、云虚拟机或容器,记作 $N_j^i$,表示归属于第 $i$ 个参与方 $P_i$ 的第 $j$ 个计算节点。

单个隐私计算平台,根据其分布式节点调度的策略差异,需要采用不同方案实现节点互联,部署相同隐私计算平台的单独业务执行方称为联盟方。节点互联互通贯穿联合计算任务全周期,在框架部署阶段,各个隐私计算框架以信息配置的方式,明确单个平台内部节点的资源、IP(互联网协议)地址和端口等信息,默认在配置完成后内部各个节点可以互信。在执行计算任务过程中,节点作为计算信息传递的载体,通过信息的互通实现计算任务的调度和有序执行成为节点互联的目标。节点互联需要依据待同步的信息类型制定不同的通信协议,如在节点通信发送 Requset 和反馈 Response 的格式、作业信息、计算信息、同步数据等内容格式、代码和对应的含义。

### 2.2.1 联盟方节点自治

每个联盟方提供一个节点或一个集群参与联合计算,每个计算节点资源调度由联盟方节点的调度功能模块完成,如图 3 所示,联合计算过程中,$P_1$ 和 $P_2$ 只开放节点 $N_4^1$ 和 $N_4^2$ 作为通信节点,传输计算过程中需要的进程信息、任务信息和数据参数等内容,相当于每个联盟方提供一个代理节点完成相关调度任务。集群管理等相关功能可能部署在代理节点上,也可能部署在单独某计算节点上,以完成多节点的管理、监控。在此种场景下,跨平台的节点互联互通是只需要面向不同框架开放代理节点即可。

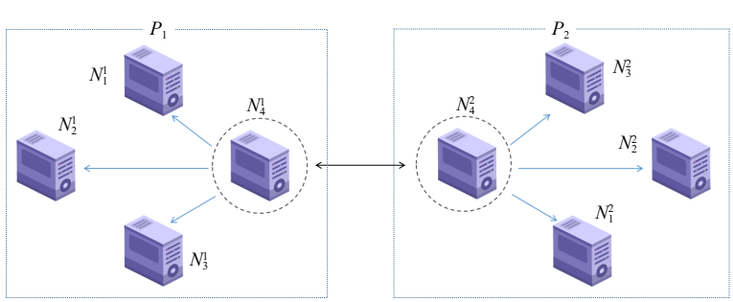


图 3 参与方节点自治场景

互联方式也有 2 种,一种是以单平台为中心,根据其开放的节点通信接口,构建协议转换模块,实现节点互联,另一种是构建独立于各平台的独立开放协议,所有平台根据协议进行面向外部协议的适配模块,实现互联互通。方式一更适合单企业主导的互联互通模式,其完善性、安全性由中心平台完成,可扩展性依托于生态规模,但可扩展性较差;方式二更适合多企业、研发方参与的联盟式生态,可扩展性强,但需要所有参与方都进行适配。

但是,值得注意的是,在每个隐私计算平台内部,其节点的信任机制由平台内部的信任机制控制,如基于证书的身份凭证验证,计算过程中基于 Token 的身份验证等,在跨平台的节点互联互通中,构建多方互信的节点认证方法,是节点互联互通后续研究的重要方向。

### 2.2.2 统一引擎管理

部分隐私计算平台采用统一管理不同联盟方计算节点的模式进行内部节点互联与调度,多采用成熟的分布式框架,如 Ray、Dask 和 Tensorflow Distributed 等框架。平台不以联盟方为区隔,采用分布式管理引擎统一管理所有计算节点,并在物理节点上虚拟形成逻辑计算设备,由逻辑计算设备调度物理计算节点,实现任务的协同计算。

在此种场景下,跨平台的节点

互联,需要连接该分布式引擎的管理节点,但是此类场景下,由于成熟分布式调度框架一般功能较为完善,其他隐私计算平台需要互联则须采用适配器模式降低改造成本,更容易形成以单平台为中心的跨平台互联模式。

### 2.3 算法互联

算法互联是业界关心和积极探索的重点。行业上,一般用“黑盒”“白盒”针对算法组件互联进行分类,白玉真等[22]根据是否为同一开发者分为跨平台算法迁移和跨开发者算法对齐。“黑盒”“白盒”的互联互通模式与参与企业是否开源其平台、主导塑造的生态模式息息相关,本节从工程化的角度,根据隐私计算平台改造分析不同算法互联的路径和特点。

算法互联的是当用户发起如隐私求交、联合统计或联合建模推理等任务时,不同隐私计算平台调用内部算法组件协同完成任务的算法执行过程互联。从互联深度区分,算法互联可以分为计算任务调度的互联和计算步骤的互联互通。计算任务调度的互联是指多方发起计算任务后,不同的隐私计算平台完成不同的 task,同步 task 的执行状态和结果数据,每个 task 调用独立的算法组件完成,同步内容与 DAG 中作业编排数据对象中的元数据结构相符。计算步骤的互联互通是不仅包含了 task 状态和数据的同步,也包含了 task 执行过程中的数据同步,如机器学习模型训练过程中的梯度更新、矩阵计算结果等。

算法互联从平台改造视角,可以分为 3 种,如表 2 所示。一是单向适配的模式,即隐私计算平台 B 开放相关调度接口,隐私计算平台 A 在平台框架内研发适配模块,将自有的数据结构、任务模型、状态参数等转换为符合 B 平台开放接口的输入类型与格式,同理,在获取 B 平台返回信息后转换为 A 平台的相应结构,实现 2 平台的互联互通。由于是单向适配,此种方法一般不涉及算法内部流程的同步,重点在于算法任务调度的互联互通,对应业界“黑盒”模式。双方需要公开的部分为隐私计算平台 B 的通信接口、通信信息的格式,隐私计算平台 A 无需部署额外插件,只需要在内部增加适配转换功能即可。典型实践如在 Fate 1.0 和 2.0 框架中,Fate 构建 OSX 通信组件,研发兼容《金融业隐私计算互联互通 API 技术文档》的通信传输接口,实现面向金融业隐私计算平台的单向适配。二是双向适配模式,双方都有适配模块一般是基于点对点的平台互联,即双方没有面向外部公开的通信协议、算法流程和步骤,在隐私计算平台发展前期,此种模式较为常见,且都是基于已成型平台进行互通,算法内部逻辑互不开放,采用算法调度层面互通。因此,也常见于“黑盒”模式。也有研发者整合 2 平台的适配模块,以独立适配器完成格式转换和数据传输等功能,但是,定制化程度高,一旦增加新平台,需要大量工程化改造。因此,独立适配器逐渐形成自有标准化格式,要求协作平台针对适配器进行调度信息、算法数据等格式转换。典型实践如百度的 HIGHFLIP(high layer federated learning inter-communication protocol)顶层通信协议,通过标准化端口、作业描述和模型,作为跨平台互联的数据结构、作业和服务的转化中间框架。这正是从双向互联到基于公开协议的多向开放互联的过渡形态。三为多向开放互联模式,任意参与方依据公开制定的协议规范,完成内部算法流程步骤的改造和算法调度的改造,即可实现算法互联。根据协议开放的算法逻辑深度,可实现浅层的算法调度互通和深层的计算逻辑互通,平台则需要完成算子改造、组件改造和协议改造等工作。上文提到的系列《隐私计算跨平台互联互通开放协议》制定了算法采用的加密算法、向量计算方法、算法流程、通信框架、消息报文格式等内容,实现了

表 2 算法互联模式

| 名称 | 改造内容 | | 适用场景 |
|---|---|---|---|
| 单向适配 | 单平台适配改造:单向算法数据格式转化、通信建立 | 基于无改造方的算法逻辑 | 主动兼容其他平台框架 |
| 双向适配 | 双平台适配改造,双方配置适配模块以转化形成对方数据格式;或采用独立适配器,完成向适配器格式转换 | 基于双方已有算法逻辑的数据格式和计算方法 | 独立小生态,承担常态化联合计算任务 |
| 多向开放互联 | 算法组件拆分封装;可插入独立算法组件适配;通信协议改造 | 基于公开协议 | 多方形成开放联合生态 |

算法层面的深度互通,并在运营商、银行、研究机构等开展实践,这是工程化成本最高、侵入性最高的方法。但是,由于面临算法开放后的知识产权问题,部分平台采用算法组件改造的模式,保障底层算法的不可见。将算法拆解成标准组件,如数据预处理、梯度计算、安全聚合等,也可将算法整体封装成独立组件,规范组件调用间的数据通信格式,保留算法底层计算方法的基础上实现算法互联。目前,行业上出现的独立第三方算法组件就是此模式的典型实践,如招商银行的隐私计算平台就采用了插拔式算法组件与其他公司平台进行算法互联互通。

### 2.4 算法互联协议

跨隐私计算平台互联互通的基础在于协议,不同的适配模式下,协议分为双方协定基础上的定制化协议和多方共识的标准化协议。从互联互通的深度上来说,算法调度互联的通信协议,主要是算法任务的调度信息的互联,算法互联则包括算法流程过程中交互的数据结构、算法内部协议的配置信息,以 ECDH-PSI 算法为例,包括算法参数协商信息、待求交的 PSI 数据量、传输的批次、批次中的密文、密文数量等信息,表 3 展示了由隐私计算联盟 2023 年发布的《隐私计算 跨平台互联互通开放协议第 1 部分:ECDH-PSI》中 ECDH-PSI 算法的参数协商数据结构。协议一般采用 gPRC 网络通信协议,采用 Prococol Buffers 进行消息的序列化和反序列化。因 gRPC 支持跨语言使用和多种消息的传输,可适配于不同的场景和平台。另外,gRPC 因其丰富的分布式治理功能,可与容器和容器的编排管理平台兼容集成,更适配于跨多域的多隐私计算平台通信。

表 3 已有规范中 ECDH-PSI 算法参数协商数据结构示例

| 属性名称 | 数据类型 | 数据说明 | 数据备注 |
|---|---|---|---|
| supported_versions | int32 list | 支持的版本列表 | 必选 |
| ec_suits | tuple〈int32,int32,int32〉list | 算法套件编号推荐列表,〈Curve编号,Hash编号,HashToCurveStrategy编号〉 | 必选 |
| point_octet_formats | int32 list | 支持的点的序列化格式的enum值,如国密SM2标准GB/T32918压缩、不压缩 | 必选 |
| supported_point_truncation | bool | 是否启用二次密文截断,true表示支持发送方截断 | 必选 |

消息队列是实现互联互通的调度组件。通过消息中间件,通过消息的产生和消费,实现互通信息的更新,可以以任务要素标识符(ID)作为消息主题,构建基于协议的消息格式,实现消息互通。

## 3 隐私计算平台互联的难点与趋势

### 3.1 互联难点

总体来看,隐私计算平台互联,仍然面临着底层技术不兼容、跨平台数据资源互操作难和复杂数据交互流程提升数据泄露风险,以及多方通信性能难以满足用户需求等难题。

从算法视角看,深层次的算法互联,技术的兼容性仍然是跨隐私计算互联平台面临的难题。如前文所述,不同隐私计算平台技术倾向的功能不一致,有的倾向于基于 MPC 的联合建模,有的倾向于基于联邦学习的联合建模,同样的任务明密文混合互通难度大,还有的平台采用区块链智能合约保障互联[22],区块链与非区块链模式的互通难。另外,各种安全和加密算法是隐私计算核心技术,不同平台采用的不同底层安全加密算法也提升了算法互联层面的难度。未来可聚焦优化协议推动底层计算语义互通,解决不同算子协议栈差异大难题,如 MPC 依赖的同态加密矩阵乘法协议、联邦学习的梯度聚合异步协议等。

从数据视角看,不同隐私计算平台采用不同数据格式和不同数据预处理方法,不同特征工程的数据处理方法也提升了跨平台互联的适配和操作复杂性,如多方安全计算采用密文,不会破坏数据分布,而联邦学习的横向联邦会导致数据的分布变化,特征工程的方法选取对训练结果会产生影响,多平台互联后对数据分布影响和训练效果难以预计。通过动态的数据分布监控,以合成数据和数据扰动

等模式降低数据分布变化的敏感性,也是解决该问题值得研究的方向。

从安全视角看,跨平台的隐私计算任务调度增加了数据流转的环节,不同技术框架和不同的数据节点都可能引入新的安全漏洞,增加数据泄露的风险。而且,跨平台的隐私计算互联,进一步增加了监管部门的监管难度,正如有些企业探索区块链和隐私计算结合的方式提升操作的可信和可溯源,但同样提高了平台的计算能耗,尤其是向 MPC 基于密文的计算,加密、解密和基于密文的计算本身消耗已经非常大,如果再结合区块链等技术,也产生了更高的能耗和效率挑战。解决能耗问题可探索在算法层优化替代密码学原语,如轻量级的同态加密、新型的 MPC 协议等都可降低计算能耗,而与区块链结合的方式,可依靠区块链的共识机制优化、链上链下分层计算、轻节点等方式进行优化。

从性能视角看,跨平台的隐私计算平台互联往往面临着跨地域的通信,基础网络能否保障跨平台的通信时延要求存在不确定性。另外,由于不同平台采用的算法和架构不一致,平台间的任务计算效率存在差异,跨平台的互联,尤其是面临大规模数据的计算任务,会带来额外的计算和通信消耗,可能会引入新的效率性能瓶颈。目前,中国正在推进光网络等网络基础设施建设,通过更高速的网络,依托“云-边-端”的分层算力调度优化,通信效率与性能瓶颈可能进一步降低。

从生态视角看,隐私计算平台互联的驱动并非仅仅是技术驱动,也来自实际场景需求的市场驱动,但是市场由大型互联网企业、科技创新型企业、研究机构和以产品盈利为目标的技术服务型公司构成,各方利益难以平衡导致隐私计算平台的互联互通难度大大提升。大型互联网企业倾向于打造以自有平台为核心的互联互通标准化协议和接口,科技创新型企业在自研隐私计算平台的基础上,依托不同的标准或实现群体性互通,或与重点交互业务方实现点对点平台互联,而技术服务型企业在各方中,通过平台架构升级和多协议适配,实现面向不同企业平台的互联适配。目前已逐渐形成了以若干平台为中心的“互联平台群岛”,部分企业倾向于构建群体的封闭生态,因此,跨行业共识的统一隐私计算开放协议仍然面临挑战。

### 3.2 发展趋势

虽然隐私计算平台互联互通面临各种挑战,但是跨平台互联仍然是未来发展的重要趋势,各个隐私计算平台主导方也在持续推动跨隐私计算平台互联互通,目前已经逐渐形成了小范围的开放互联实践。

标准和规范方面,金融行业、电信行业都在探索行业内部的隐私计算平台互联标准。未来如何在符合不同行业监管要求和业务需求的基础上统一各行业标准、形成可兼容的标准体系成为了隐私计算平台各研发方的重点研究方向。目前来看,跨行业标准的适配,如通信行业在推进跨隐私计算平台互联互通过程中,开始兼容金融行业的标准。另外,随着国内数据要素市场的繁荣发展,跨隐私计算平台互联的场景需求将持续提升,标准快速落地工程化难度和可规模化复制性也成为了制定标准需要重要考量的问题之一。

技术方面,双向适配模式下的双方适配功能将逐渐被“适配器”式的转换产品替代,而规模化的、基于开放协议的平台互联互通将逐渐成为更多技术方的选择。多技术融合可能成为隐私计算跨平台互联互通完成计算任务的辅助手段,性能提升和安全保障也将成为因计算技术互联互通的重要研究方向。但是,基础算子的互联互通面临的挑战比较高,目前只能实现个别算法的基础算子互联互通。如何保障各方知识产权、利益分配合理基础上,实现各算法的进一步互联互通也成为后续重点研究问题。多方采用统一算法网关和逻辑接口,在算法网关内实现算法注册和路由,由平台内部完成算法“黑盒”实现,接口算法调用和算法实现解耦,也可能成为解决这个问题的方案之一。另外,在隐私计算互联互通方案中,关于可信执行环境的互联探索和实践相对较少,主要是因为可信执行环境需要特定的硬件设备,本身应用范围相对较窄,如上文提到,北京金融科技产业联盟提出了基于远程证明报告的互通,后续如何提升报告中心节点的安全性保障和拓展其他方式的可信执行环境互联也是重要的研究方向。

总体来说,隐私计算平台作为数据流通基础设施的重要组成,对互通方案的可落地性要求会随着应用场景的增加而更加严格,各个

参与主体由于商业考量导致生态小范围内闭源和技术差异导致的互通难度大都将随着市场需求的增加和技术进步逐渐实现突破。

## 4 结论

梳理了相关文献和公开资料,从功能视角切入当前隐私计算平台,结合隐私计算平台互联互通实践,总结了当前主要的互联方案。目前已形成以金融行业和互联网企业为中心的小规模探索型互联互通生态。同时结合当前生态,根据隐私计算平台的算法互联,从落地难度视角出发,明确隐私计算平台互联互通的适配难度、方式,从适配的工程化、复杂性角度阐述各种模式与当前行业“黑盒”“白盒”式互联的关系。

隐私计算的互联互通仍然处在小范围的试点实践过程中,单场景的应用较多,规范化的落地实践仍然面临着技术、监管和生态上的挑战。但是,随着鼓励数据开放、数据要素流通的政策文件频发,中国不断增加的数据流通需求将有效驱动技术的进步,而生态开放将会成为制约隐私计算互联互通的核心要素之一。因此,构建兼容并存、大规模公开互联、小规模封闭互联的模式可能成为未来隐私计算互联互通的重要模式。

### 参考文献(References)


[1] Prasser F, Eicher J, Spengler H, et al. Flexible data anonymization using ARX: Current status and challenges ahead[J]. Software: Practice and Experience, 2020, 50(7): 1277−1304.

[2] Cape Privacy. PyCape: The cape privacy python SDK[EB/OL]. (2023−03−31) [2024−08−01]. https://github.com/capeprivacy/pycape.

[3] Holohan N, Braghin S, Mac Aonghusa, et al. Diffprivlib: The IBM differential privacy library[EB/OL]. (2019−07−04) [2024−08−01]. https://arxiv.org/abs/1907.02444.

[4] Liu C, Wang X S, Nayak K, et al. ObliVM: A programming framework for secure computation[C]//Proceedings of IEEE Symposium on Security and Privacy. Piscataway, NJ: IEEE, 2015: 359−376.

[5] Inpher. Inpher XOR platform unlocking value and secure data collaboration through privacy−preserving nachine learning and analytics[R]. Switzerland: Inpher, 2023.

[6] UC Berkeley RISE Lab. PyCape: A platform for secure analytics and machine learning[EB/OL]. (2021−06−18) [2024−08−01]. https://github.com/mc2-project/mc2.

[7] Microsoft. Microsoft SEAL is an easy−to−use and powerful homomorphic encryption library[EB/OL]. (2023−01−11) [2024−08−01]. https://github.com/Microsoft/SEAL.

[8] Lepoint T, Patel S, Raykova M, et al. Private join and compute from PIR with default[M]//Advances in Cryptology: ASIACRYPT 2021. Cham: Springer International Publishing, 2021: 605−634.

[9] Ziller A, Trask A, Lopardo A, et al. PySyft: A library for easy federated learning[M]//Federated Learning Systems. Cham: Springer International Publishing, 2021: 111−139.

[10] Bogdanov D, Laur S, Willemson J. Sharemind: A framework for fast privacy−preserving computations [M]//Computer Security: ESORICS 2008. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008: 192−206.

[11] DropoutLabs. A framework for encrypted machine learning in tensorFlow[EB/OL]. (2023−02−08) [2024−08−01]. https://github.com/Microsoft/SEAL.

[12] Google. An open−source framework for machine learning and other computations on decentralized data [EB/OL]. (2024−7−29) [2024−08−01]. https://github.com/OpenFederatedLearning/federated.

[13] Knott B, Venkataraman S, Hannun A, et al. Crypten: Secure multi−party computation meets machine learning[J]. Advances in Neural Information Processing Systems, 2021, 34: 4961−4973.

[14] WeBank. An industrial grade federated learning framework[EB/OL]. (2023−12−21) [2024−08−01]. https://github.com/FederatedAI/FATE.

[15] Ant Group. A unified framework for privacy−preserving data analysis and machine learning[EB/OL]. (2024−06−25) [2024−08−01]. https://github.com/secretflow/secretflow.

[16] 北京金融科技产业联盟. 金融业隐私计算互联互通技术研究报告[R]. 北京: 北京金融科技产业联盟, 2023.

[17] 招商银行隐私计算互联互通项目组. 创新隐私计算技术, 实现异构平台互联互通[EB/OL]. (2022−02−10) [2024−08−01]. https://www.secrss.com/articles/39097.

[18] 微众银行. FATE 2.0 版本互联互通接入指南[EB/OL]. (2023−12−31) [2024−08−01]. https://federatedai.github.io/FATE-Flow/latest/zh/fate_access.

[19] 隐私计算联盟. 隐私计算平台互联互通文档[EB/OL]. (2023−07−13) [2024−08−01]. https://interconnection.readthedocs.io/zh-cn/latest.

[20] 隐私计算联盟. 隐私计算白皮书(2022年)[R]. 北京: 隐私计算联盟, 2022.

[21] 叶剑, 李文. 支持互联互通的隐私计算网关设计与实现[J]. 大数据, 2023, 9(6): 28−38.
[22] 白玉真, 杨靖世, 袁博. 隐私计算互联互通发展现状与展望[J]. 信息通信技术与政策, 2024(5): 71−78.

# The interconnection path and challenges of privacy computing platforms in the era of data elements

LIU Yinghui[1,2], WEI jinwu[2*], ZHANG rongfang[2], CAI yixin[2], LI kun[2]

1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
2. China Unicom Research Institute, Beijing 100176, China

**Abstract** With the confirmation of data elements as the latest production factor, accelerating data sharing and value mining on the basis of ensuring data security has become an industry consensus. Privacy computing is the core technology for achieving "data available but invisible". Privacy computing platforms have flourished and gradually forming a phenomenon of "platform islands" with privacy computing platforms as the core. How to solve platform silos and further promote the integration and value of data elements has become an important research direction. This paper summarizes the typical functions of privacy computing platforms at home and abroad, the current status of the interconnection ecology of privacy computing platforms in China, and industry practice cases. The research scope of the paper is clarified as node interconnection and algorithm interconnection. It analyzes the different modes of node interconnection as well as algorithm, and finally summarizes the algorithm, data, security, and performance difficulties of privacy computing cross platform interconnection, and elaborats on the trend and prospects. Although privacy computing platforms face various challenges in interconnection, cross−platform interconnection remains a significant trend for future development. In terms of standards and regulations, both the financial industry and the telecommunications industry are exploring their own internal standards for interconnecting privacy computing platforms. From a technical perspective, the dual adaptation functions under the two−way adaptation model will gradually be replaced by "adapter"−style conversion products. Meanwhile, large−scale, open protocol−based platform interconnection is increasingly becoming the preferred choice for more technical providers.

**Keywords** data elements; privacy computing; node interconnection; algorithm interconnection

（责任编辑 王微）