



北京金融科技产业联盟
BEIJING FINTECH INDUSTRY ALLIANCE

隐私计算技术金融应用 研究报告

北京金融科技产业联盟

2022 年 2 月

版权声明

本报告版权属于北京金融科技产业联盟，并受法律保护。转载、编摘或利用其他方式使用本报告文字或观点的，应注明来源。违反上述声明者，将被追究相关法律责任。



编制委员会

主 编：

潘润红

编委会成员：

何 军 左 燕 赵焕芳 聂丽琴

编写组成员：

杜 俊	耿 博	蒙永明	黄本涛	司忠平	徐安滢
樊明璐	滕 竹	强 锋	魏博言	王光中	傅 杰
葛明嵩	焦惠芸	何东杰	彭 晋	殷 山	袁鹏程
昌文婷	赵 伟	郝 洁	王 榕	杨少杰	孙丽洁
王 雪	李武璐	霍昱光	范 涛	刘筠璨	吴博峰
詹明星	王跃东	窦永金	姚 明	李 博	何 浩
王湾湾	郑华祥	刘 姝	龚自洪	郭 超	王 超
李克鹏	苏庆慧	刘 江	王煜惠	王铀之	倪裕芳
戎艳中	高 扬	马文婷	王 鹏	贾雪丽	樊昕晔
李 钰	陈嘉俊	张敬之	曹旭涛	林冠峰	屈 阳
孙赞美	陈 浩	袁 晔	黄小刚	彭宇翔	张明明
孟 丹	卞 阳	黄翠婷	方 竞	金银玉	单进勇
蔡超超	王云河	靳 晨	时 代	林佳萍	张 煜
高志民	王健宗	黄章成	卢春曦	李泽远	邱晓慧
杨 波	陈 鑫	张晓武	胡祎波	周 辉	董 蔚

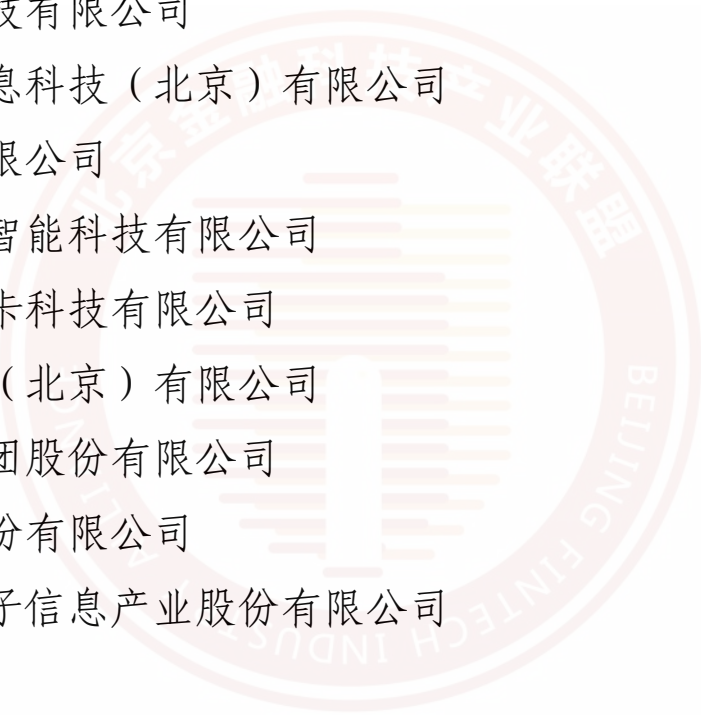
主 审：黄本涛 司忠平

统 稿：郭 栋 刘宝龙

参编单位：

北京金融科技产业联盟秘书处
中国农业银行股份有限公司
成方金融信息技术服务有限公司
中国工商银行股份有限公司
交通银行股份有限公司
招商银行股份有限公司
中国银联股份有限公司
蚂蚁科技集团股份有限公司
建信金融科技有限责任公司
深圳前海微众银行股份有限公司
中国银行股份有限公司
北京国家金融科技认证中心有限公司
中国信息通信研究院
深圳市洞见智慧科技有限公司
矩阵元技术（深圳）有限公司
蓝象智联（杭州）科技有限公司
腾讯云计算（北京）有限责任公司
北京瑞莱智慧科技有限公司
光之树（北京）科技有限公司
上海浦东发展银行股份有限公司
光大科技有限公司

浙商银行股份有限公司
中国民生银行股份有限公司
百度在线网络技术（北京）有限公司
北京融数联智科技有限公司
同盾科技有限公司
上海富数科技有限公司
北京数牍科技有限公司
华控清交信息科技（北京）有限公司
华为技术有限公司
深圳壹账通智能科技有限公司
北京银联金卡科技有限公司
度小满科技（北京）有限公司
云从科技集团股份有限公司
平安银行股份有限公司
云南南天电子信息产业股份有限公司



目 录

一、概述	1
二、隐私计算相关技术分析.....	3
（一）多方安全计算	3
（二）联邦学习	16
（三）可信执行环境	21
（四）同态加密	23
（五）差分隐私	25
（六）零知识证明	26
（七）数据去标识化及脱敏	26
（八）相关技术结合	30
（九）技术对比分析	40
三、隐私计算金融应用背景.....	44
（一）政策与法律法规	44
（二）国内外相关标准	47
四、隐私计算金融应用场景.....	50
（一）智能风控场景	50
（二）智能营销场景	72
（三）智能运营场景	84
（四）隐私信息检索场景	88
（五）供应链金融场景	89
（六）反洗钱场景	91
（七）企业级数据流通交易	95
（八）平台应用	107
五、问题与建议	110
（一）风险与挑战	110
（二）发展建议	114
附录：隐私计算技术平台产品.....	121

一、概述

2021 年 12 月，中国人民银行发布《金融科技发展规划（2022—2025 年）》（银发〔2021〕335 号文印发）明确提出从强化数据能力建设、推动数据有序共享、深化数据综合应用、做好数据安全保障方面充分释放数据要素潜能，并陆续出台了金融数据安全相关标准。金融数据安全与共享应用的重视程度与日俱增。

为推动金融业务更好地开展，推动跨机构、跨地域、跨行业数据资源有序共享，实现数据可用不可见、数据不动价值动，隐私计算技术被重点关注和推广。隐私计算技术在金融行业的实践中，已逐步实现产品化、平台化，通过规模化推广，能够在金融领域的更多业务场景、更全面的上下游供应链体系应用中，更好发挥数据价值，避免数据滥用，并产生极好的经济效益。在数据成为生产要素并推定流通的情况下，隐私计算技术及在行业的应用正在加速发展。金融机构和科技公司纷纷开始建设隐私计算平台，在精准营销、信贷风控、信息共享、反洗钱等领域进行试点。

本报告通过对隐私计算技术的研究及应用场景的探索，能够对隐私计算在金融行业的应用提供参考和指导。对隐私计算多方合作模式的研究，能够促进不同机构、企业在符合各项法律、法规及政策的前提下，进行顺畅高效的数据合作，解决“信息隐私”和“数据孤岛”问题，达成合作共赢。本报告围绕隐私计算，展开政策法规、标准、技术、场景调研，对应用场景进行探索实验，

形成解决方案，并发布技术研究报告。本报告一共分为五章，组织结构如下：

第一章是整体概况，介绍了本报告的研究内容与意义。

第二章介绍了隐私计算的关键技术、隐私计算与其他前沿技术的结合以及隐私计算技术的对比总结。

第三章是金融行业隐私计算背景，介绍了与隐私计算相关的政策、法律法规及标准。

第四章列举了金融行业隐私计算的应用场景与案例实践，涉及风控、营销、运营、匿踪查询、供应链金融、反洗钱等诸多业务领域。

第五章分析了金融行业推进隐私计算发展所面临的风险与挑战，并从行业政策、标准化、技术发展和产业发展的角度提出了相关建议。

二、隐私计算相关技术分析

金融业应用的隐私计算核心技术包括多方安全计算、联邦学习、可信执行环境、同态加密、差分隐私、去标识计算及脱敏技术等，本章介绍这些技术的原理实现方案、难点与创新以及发展方向等^[1]。

（一）多方安全计算

1. 总体介绍

多方安全计算（Secure Multi-Party Computation，简称 MPC）指在分布式网络环境下，不依赖可信第三方代替各参与方进行计算，而是由各对等的参与方通过网络协同共同完成某一计算任务。通常情况下，由两个或多个持有私有输入的参与方，在不泄漏各自私有输入信息的情况下联合计算一个函数，各自得到他们预定的输出。

MPC 的基本计算模型如图 2-1 所示，在分布式协同计算网络中， N 个互不信任的参与者集合，各自拥有秘密数据。 N 个参与者协同计算函数，其中， Y_i 是第 i 个参与者需要获得的结果。在此过程中，任何参与者不能获得除了 Y_i 之外的其他秘密信息。

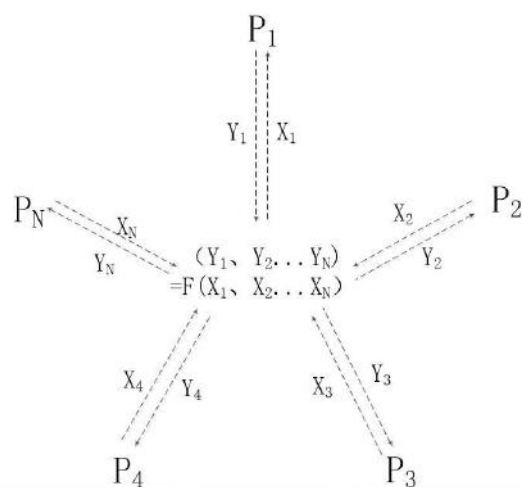


图2-1 MPC基本模型

多方安全计算通常考虑如下框架性的安全要求。其中正确性和隐私性是必须满足的安全要求。

正确性 (Correctness)

MPC 系统中每个参与方获得的输出都是正确的。

隐私性 (Privacy)

MPC 系统中每个参与方除了各自的预期输出之外，无法获得其他额外的信息。换言之，当前参与方唯一可获得的关于其他参与方的信息，仅能从当前参与方的输出结果中推导。

输入独立性 (Independence of Inputs)

任意的参与方无法通过其他参与方的输入来构建自己的计算输入。

输出保证 (Guaranteed Output Delivery)

不诚实的参与方无法阻止其他诚实的参与方获得输出。换言之，不诚实的参与方无法通过发动类似“拒绝服务攻击”来阻断

多方安全计算协议。

公平性 (Fairness)

当且仅当诚实的参与方获得计算输出的时候，不诚实的参与方才能获得计算输出。

从安全性而言，MPC 通常考虑协议中存在攻击者，该攻击者可以控制部分参与方，在协议运行过程中产生不诚实的行为。MPC 设计建模过程中，通常考虑如下关于攻击者能力的安全假设：

半诚实敌手 (Semi-honest Adversaries)

MPC 安全假设的一种，如果 MPC 系统中的不诚实参与方会按照协议规定进行运算，但试图通过协议中得到的信息挖掘其他参与方的隐私，MPC 系统除计算结果外不会泄露任何参与实体的隐私。

恶意敌手 (Malicious Adversaries)

MPC 安全假设的一种，如果 MPC 系统中的不诚实参与方试图通过改变协议行为来挖掘其他参与实体的隐私，MPC 系统除计算结果外不会泄露任何参与方的隐私。

隐蔽敌手 (Covert Adversaries)

MPC 安全假设的一种，如果 MPC 系统中的不诚实参与方试图破坏协议执行，MPC 系统有能力检测到该攻击行为并作出相应的处罚。

2. 基础技术

从 MPC 提出到现在,学术界和产业界针对各种多方安全计算模型设计了多方安全计算协议,这些协议广泛的使用了秘密分享、不经意传输和混淆电路等基础密码学技术。

(1) 秘密分享

秘密分享概念最早由密码学家 A. Shamir 和 G. Blakley 于 1979 年分别独立提出。秘密分享 (secret sharing, SS) 也被称为秘密分割,是一种对秘密信息的管理方式,将秘密进行拆分,拆分后的每一个子秘密由不同的参与者管理,单个参与者无法恢复秘密信息,需要多个参与者一同协作进行合并才能恢复秘密。秘密共享的应用十分丰富,数字签名、电子拍卖、电子选举是典型的应用场景。秘密分享自身是保护数据安全的重要手段,同时也是多方安全计算的基础技术之一。

典型的秘密分享方案由秘密分割算法和秘密重构算法两部分组成。秘密分割算法将主秘密分割成多份子秘密。秘密重构算法通过一组授权的参与者的子秘密恢复主秘密,如图 2-2 所示。

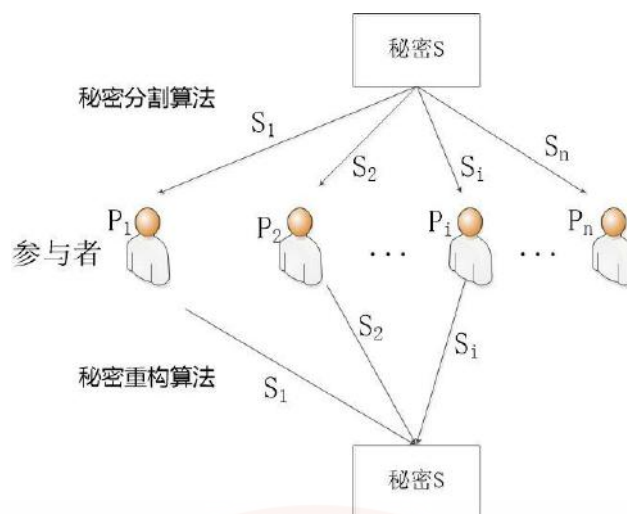


图2-2 秘密分享方案

为满足不同应用场景的使用需求，秘密共享衍生出了可验证秘密共享、动态秘密共享、多秘密共享等理论方向。最简单的秘密共享方案是 Shamir 的门限秘密共享方案，利用拉格朗日插值定理，实现了门限秘密分享，有两个参数： n 和 t ， n 表示参与分割秘密信息的参与者的数量， t 表示至少几个参与者聚到一起才可以恢复出秘密信息。该方案原理描述如下：

将主秘密 S 拆分成 n 份，每个参与者获得一份，每一份被称作一个子秘密，每个参与者秘密地保存子秘密。

拆分时，预先设定至少 t ($t \leq n$) 个参与者聚到一起才可以恢复主秘密 S 。

需要恢复主秘密 S 时，有至少 t 个参与者，通过计算恢复出主秘密 S 。

秘密共享方案必须至少满足可重构性和安全性两个性质。具体如下：

可重构性, 任何 t 个或者多于 t 个的参与者的子秘密放在一起, 通过秘密重构算法可以恢复出主秘密 S 。另一方面, 即使 $n-t$ 个参与者遗失了子秘密, 剩下的 t 个参与者仍然可以恢复出主秘密^[2]。

安全性, 已知少于 t 个子秘密的信息时, 不能恢复出主秘密 S , 即少于 t 个参与者即使合谋也得不到主秘密 S 。

针对 Shamir 可能存在密钥持有者作恶的问题, 改进的协议如下:

Feldman 的 VSS 协议: 在常用的秘密分享方案基础上附加验证操作, 实现可验证的秘密分享。密钥持有者发送密钥的时候, 还要附加一个承诺, 这一承诺是密钥持有者拥有该密钥的一个证据, 密钥接收者接收到密钥后, 首先验证承诺, 验证通过后, 再接受密钥碎片。

(2) 不经意传输

不经意传输 (Oblivious Transfer, OT) 是一种可保护各方隐私的秘密选择协议, 是设计多方安全计算协议时经常使用的一个基础工具。不经意传输协议中, 消息接收者根据一个选择比特, 从消息发送者那里获得对应的消息。从隐私保护的角度来看, 消息接收者仅能获得其选择的那条消息, 对于剩余的其他消息无从所知; 消息发送者仅能知晓其提供了一次服务, 但是无法获知哪一条消息被接收。

不经意传输 (oblivious transfer) 最初由 O. Rabin 在 1981 年提出。1985 年, Shimon Even 提出了更加实用的 2 选 1 不经意传输 (1 out 2 oblivious transfer) 协议, 如图 2-3 所示。该协议中, Alice 拥有两条秘密信息 (m_1 、 m_2), Bob 提供一个输入 (0 或者 1), 并根据输入获得输出信息, 在协议结束后, Bob 得到了自己想要的那条信息 (输入 i , 则得到 m_i), 而 Alice 并不知道 Bob 最终得到的是哪条秘密信息。后来, 2 选 1 不经意传输被扩展到 N 选 1 不经意传输协议、 N 选 K 不经意传输协议等 [3]。



图2-3 2选1不经意传输

当前, 不经意传输协议主要基于数论中的困难性假设, 使用公钥密码学算法构造。为了提高协议的效率, 主要有两种协议设计思路: 第一种设计思路是基于不同的数学理论构造更加高效的茫然传输协议; 第二种设计思路使用茫然传输扩展技术 (OT Extension), 该设计思路和基于混合加密技术的数字信封技术十分相似。相比于对称密码学操作, 基于数论中各种困难性假设的公钥密码学操作是更加昂贵的, 即算法的复杂度更高。因此, 减少公钥密码学操作, 通过使用高效的对称密码学操作取代部分

公钥密码学操作来实现不经意传输协议似乎是一个不错的选择。基于这种设计思路，不经意传输扩展技术首先使用公钥密码学操作产生少量种子 OT，然后利用对称密码学操作（如哈希函数等）将这些种子 OT 协议扩展为任意数量的 OT 协议。

不经意传输协议必须至少满足正确性、发送方隐私性、接收方隐私性三条性质：

正确性

协议执行完毕后，接收方能够获得他所选择的 k 个信息。

发送方隐私性

协议执行完毕后，接收方通过发送方发来的 n 个秘密数据，已经获得的 k 个消息以及系统的公共参数，不能获得自己没有选择的 $n-k$ 个消息中的任何一个。

接收方隐私性

协议执行完毕后，发送方通过系统的公共参数，接收方发来的选择数据，不能区分出接收方选择了哪 k 个消息。

列举部分不经意传输协议如下：

Naor-Pinkas 协议：基于离散对数困难问题，实现了二选一不经意传输。通过三次公钥密码学操作实现了半诚实模型下的 1-out-of-2 不经意传输协议。

Tzeng (2004) 协议：基于 DDH (Decision Diffie-Hellman) 假设的 n 选 1 不经意传输协议。

KKRT (2016) 协议：不经意传输扩展协议，多 OT 运行实例的并

行优化协议。

(3) 混淆电路

混淆电路 (Garbled Circuit, GC) 思想是利用计算机模拟集成电路的方式来实现安全计算, 将计算任务转化为门电路的形式, 并且对每一条线路进行加密, 在很大程度上保障了参与者的隐私安全。混淆电路 (Garbled Circuits) 的思想起源于图灵奖获得者姚期智, 因此也被称为“姚氏电路”。

混淆电路基本原理如下: 发送方将计算逻辑转化为布尔电路, 针对电路中每个门进行加密处理。然后, 发送方将加密电路 (即计算逻辑) 和加密后的标签输入给下一个参与方。接收方通过不经意传输按照输入选取标签, 对加密电路进行解密获取计算结果。

典型的混淆电路方案包括两个阶段: 第一阶段电路生成阶段和第二阶段电路执行阶段。电路生成阶段将安全计算函数转化为布尔电路。电路执行阶段则通过使用 OT、加密等方法执行电路。

第一阶段, 电路生成者首先以安全计算为目的, 生成布尔电路。然后根据双方的输入值范围和中间输出结果范围随机生成映射标签, 然后将这些标签作为密钥, 通过分组密码对每个门电路对应的输出真值表进行加密, 并对加密后的真值表值进行加扰, 从而获得混淆电路。

第二阶段, 电路生成者在执行阶段需要将自身输入对应的标

签随同混淆电路一起发送给电路执行者；电路执行者根据自己的输入信息通过不经意传输的方式选择自己对应的标签，这样电路发生者和电路执行者都无法得到对方的输入信息；电路执行者获取双方输入的对应标签后，将其作为解密密钥，对输出真值表进行解密，最终获得对应的门电路输出值。

常见混淆电路方案如下：

姚氏估值方案：使用混淆真值表，构造布尔电路，通过电路计算来实现多方安全计算的目标函数的计算。

GMW 协议：支持多方的半诚实的安全计算协议，计算和通信效率高。与姚氏混淆电路方案的不同之处在于，GMW 混淆电路方案不需要使用混淆真值表，因此无需一些查表和加解密操作，节省了非常大的计算量和通信量。

BGW 协议：利用门限秘密分享 Shamir 协议的同态性实现混淆电路，阈值 k 可根据安全性要求调整。

常见混淆电路方案对比如表 2-1 所示。

表 2-1 混淆电路方案对比

方案名称	姚氏方案	GMW 协议	BGW 协议
实现方式	混淆真值表	秘密分享与 OT	门限秘密分享 Shamir 协议
支持的参与方数量	两方	多方	多方
协议的执行轮数	常数	电路深度	电路深度
支持的电路类型	布尔电路	布尔电路或算数电路	布尔电路或算数电路

3. 特定技术

特定 MPC 包括隐私求交和隐私信息检索。

(1) 隐私求交

隐私求交 (Private Set Intersection, PSI) 是指参与方使用各自的数据集计算交集, 但不泄露交集以外的任何数据。

保护集合的隐私性在很多场景下是自然甚至是必要的需求, 比如当集合是某类用户的通讯录, 这样的输入就一定要通过密码学的手段进行保护。根据使用的场景不同, PSI 协议有以下五条技术路线^[4]:

(a) 基于公钥密码学的 PSI 协议

基于公钥密码学的 PSI 协议, 主要的用途是隐私保护场景下选择偏好的匹配, 两方可以在保护各自输入私密性的前提下验证各自的输入是有某些程度上的匹配。基于 Diffie-Hellman (DH) 框架的 PSI 协议底层可以使用基于 RSA 公钥体系来求交集, 也可以使用运行速度更快的椭圆曲线, 两种方式计算和通信复杂度都是随着集合大小线性增长^[5]。

(b) 基于不经意传输的 PSI 协议

可以实现吞吐量大性能良好的隐私求交平台, 并通过求交过程的流程控制, 防止恶意方多次暴力撞库的问题出现。不经意传输扩展协议的出现使得 OT 协议的性能有了极大的提升, 许多基于 OT 扩展协议的 PSI 达到了与哈希协议复杂度在同一量级的性

能。使用布隆过滤器和 OT 扩展协议的 PSI 协议，构造出的协议拥有可以在亿级规模的集合上操作的能力，并且可以从理论上证明在半诚实模型和恶意模型下是安全的。

(c) 基于电路的 PSI 协议

PSI 协议是解决多方安全计算的问题，通过电路来实现计算交集的函数，具有通用性和易扩展性。将计算交集的电路的输出连到计算其他电路的输入可以计算集合大小、集合中元素的求和等函数，全部过程都满足保护输入的隐私性。这种易于扩展的性质是其他基于公钥加密或是基于不经意传输扩展协议的 PSI 所没有的，因此这也是一种有实际应用价值的 PSI 方案。

(d) 基于全同态加密协议的 PSI 协议

全同态加密是一种功能强大的加密原语，允许运算电路直接在密文上计算，而不必首先解密数据。该协议拥有很小的通信量方面的代价，尽管使用了 batching, windowing, modulus switching 等优化方法，计算开销庞大仍然是未来需要解决的首要问题。

(e) 基于可信执行环境的 PSI 协议

传统的 PSI 是基于密码难题来构建可信根的。另外一种构建可信根的技术是使用可信硬件，可信根比其他的可信根（比如 TPM）都要更小更可信，可以提供一个安全可信的运行环境，保障代码和数据的机密性和完整度，并且可提供近似原生态的运行速度，从而大大降低运行时期的性能开销。

(2) 隐私信息检索

隐私信息检索 (PIR)，是指客户端检索数据库的一种方法，且数据库无法知道客户端检索的具体信息。PIR 最早是 1995 年由 Chor 等人提出，他们的方案严重依赖于通信效率，特别在海量数据检索应用场景中难以实际落地。后来理论和实践界研发了多种方案突破 PIR 的性能限制，主要分为三类：

(a) 信息论安全的隐私信息检索

此类的 PIR 提供一个很强的安全概念，在假设攻击者的计算能力是无限制的条件下保证用户的隐私性能够完全的被保护。信息论安全的隐私信息检索使用分布式数据库，每个数据库存储消息的一部分分片，而数据库之间假定是无法合谋的。

(b) 计算安全的隐私信息检索 (cPIR)

假设攻击者的计算能力在多项式的时间内是有限制的，常基于一些密码学上的计算难题，具有一定的实用意义。cPIR 的发展路线主要有四个方向，一是结合同态加密技术，将具有同态性质的 Paillier 加密方法应用于构造单数据库的隐私信息检索协议，对于半诚实或者恶意的发送方模型都是安全的。二是使用秘密分享的思路，这个方案有一个限制条件就是需要服务端数量大于 1，通过非合谋的多个服务器共同持有秘密分片隐私信息检索。三是与不经意传输结合，优化计算速度，服务端也无法指导用户最终查询信息，此方案缺点是增加了通信传输。四是借用一些处理技巧优化查询速度，如布隆过滤器，布谷鸟哈希函数，局部可

解码编码技术等。

(c) 基于硬件的隐私信息检索

基于安全硬件的 PIR 协议，借助一些安全硬件作为辅助设施，PIR 具有很高的执行效率，安全性主要取决于硬件的安全性。

按参与方是否诚实，隐私信息检索还可划分为恶意模型的隐私信息检索和半诚实模型的隐私信息检索。按照有无服务器辅助分类可分为无服务器辅助的 PIR 和有服务器辅助的 PIR。

(二) 联邦学习

1. 总体介绍

联邦学习 (Federated Learning, FL) 也称为隐私保护机器学习，是目前产学研各界都高度重视的多方协同训练机器学习模型的新范式，可以保证多方的数据不离开自己定义的安全域，是一种只通过传输中间结果（一般为学习模型的梯度信息）进行信息交换、模型聚合的联合训练机器学习模型方法。允许数据保留在各自的本地不出库，极大程度解决了一些组织的数据安全顾虑，实现了数据“可用不可见”的安全流通。谷歌在 2016 年成功地应用联邦学习，在不直接获取用户输入信息、保障用户隐私的情况下，实现了谷歌输入法的智能预测功能^[6]。

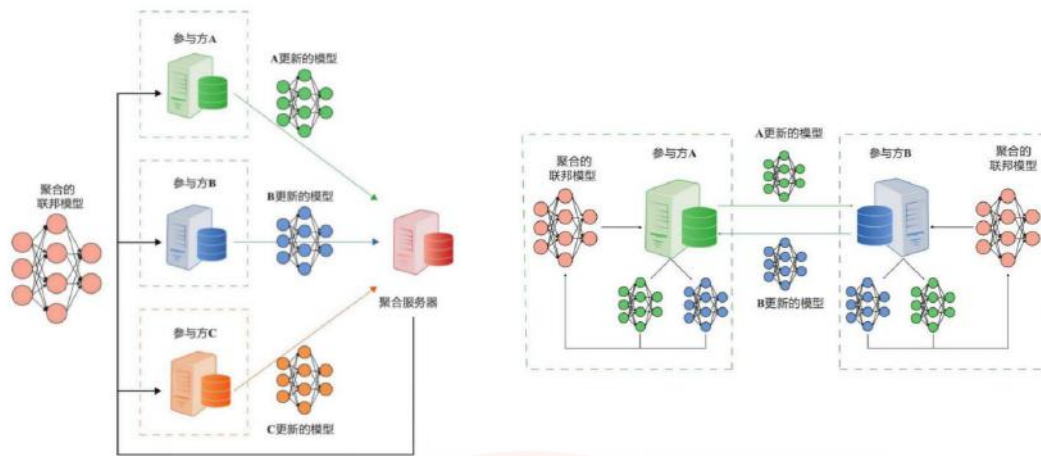
一般情况下，训练一个机器学习模型，需要先准备好所有的数据。如果数据来自于不同的所有者，需要把数据汇集起来，让机器进行模型训练。

2. 联邦学习架构

按照架构分类，联邦学习可以分为有中心服务器的联邦学习（服务器-客户端架构）和无中心服务器的联邦学习（对等网络架构）。

最初始的联邦学习是中心服务器-客户端的架构。各个客户端是各个独立的数据拥有方和模型训练方，中心服务器是模型分发方和进行安全聚合计算的一方。大致的流程如下：中心服务器向各个客户端发送具体任务的模型参数；各个客户端收到模型后，在本地利用自己的数据进行训练，并把模型参数发送给中心服务器；中心服务器收到各方的模型参数，并对其进行安全聚合运算，完成一轮模型参数迭代；如此进行多轮迭代，最终获得一个优质的全局模型。有中心服务器的流程如图 2-4（a）。前面提到的谷歌输入法的例子就是有中心的联邦学习架构。

无中心服务器的联邦学习的安全聚合是由各客户端之间通过通信交换参数，自行进行聚合运算得出的，如图 2-4（b）。由于对等网络架构中不存在中心服务器，训练方们必须提前商定发送和接收模型参数信息的顺序。



(a) 客户端-服务器架构的联邦学习 (b) 对等网络（点对点）架构的联邦学习

图2-4 两种联邦学习架构

客户端-服务器架构，协调方是一台聚合服务器，在联邦学习过程中负责模型参数的“分发”和“聚合”；对等网络（点对点）架构，不需要协调方，因为各方无须借助第三方便可以直接通信，但是会消耗更多的计算资源对消息内容进行加密和解密。

3. 联邦学习类型

按照合作的数据特点分类，联邦学习可以分为纵向联邦学习、横向联邦学习和联邦迁移学习。通常，数据矩阵的横向表示不同的数据样本，纵向表示不同的数据特征（或者标签）。一行表示一条训练样本。

(a) 横向联邦学习

横向联邦学习是通过增加样本的数量的方式获得高质量模型的方法。这些数据合作方的数据具有特征重叠较多、样本重叠

较少的特点。横向联邦学习较适用于同行业间的合作。例如，某几家银行都各自拥有有限数量的类似业务的客户，通过横向联邦学习，可以达到保护客户隐私的情况下，聚合所有客户训练一个样本数量足够多的客户风险预测模型。横向联邦学习数据合作示例如图 2-5 所示。

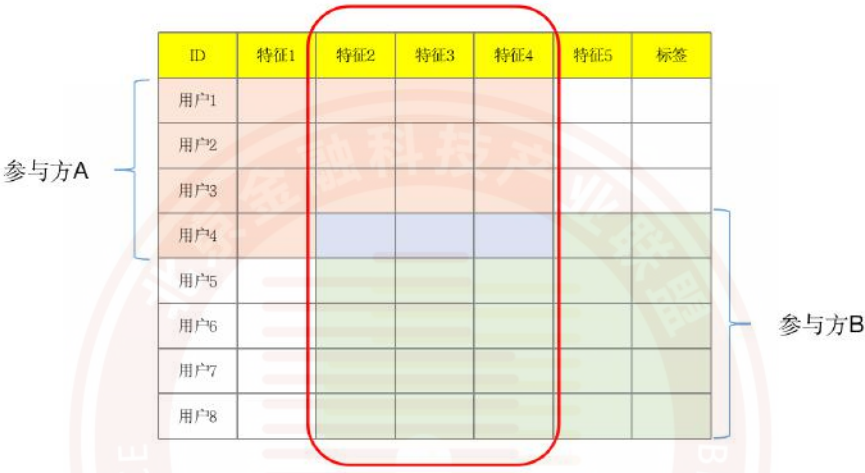


图2-5 横向联邦学习数据合作示例

(b) 纵向联邦学习

纵向联邦学习是通过增加样本特征维度的方式获得高质量模型的方法。这些数据合作方具有样本重叠较多、特征重叠较少或不重叠的特点。纵向联邦学习较适用于跨行业间的合作。例如，银行的贷中监测模型中，银行只有用户的基本信息和银行账户信息，而如果加入用户的运营商的通话行为数据或者更加丰富的用户画像数据，就可以以更加多元化地角度去评估该用户的风险指数。纵向联邦学习数据合作示例如图 2-6 所示。

	ID	特征1	特征2	特征3	特征4	特征5	标签
参与方A	用户1						
	用户2						
	用户3						
	用户4						
	用户5						
	用户6						
参与方B	用户7						
	用户8						

图2-6 纵向联邦学习数据合作示例

(c) 联邦迁移学习

联邦迁移学习是应用于样本和特征均重叠较少的情况下，利用迁移学习技术来学习不同数据集的“知识”的方法。例如，有两家不同的机构，一家是位于中国的银行，另一家是位于美国的电商。由于受地域的限制，两家机构的用户群体交集很小。同时，由于机构类型的不同，二者的数据特征重叠也较少。这种情况下，我们不对数据进行切分，利用迁移学习来解决单边数据规模小的问题。

联邦学习是目前受到全世界广泛关注的重要技术，将会在以数据为要素的市场环境中，发挥其重要的作用。与此同时，联邦学习因为其在传输中间结果的过程中，有可能被恶意参与者进行模型投毒，或者恶意合作方会通过中间结果进行反推出隐私信息

的风险，所以，如何攻克这些问题也成为另一个探索方向。目前常用的解决方法为使用同态加密、差分隐私或多方安全计算等技术，对参数信息进行加工，并将加工后的数据（加密梯度）发送给聚合服务器。

（三）可信执行环境

可信执行环境（Trusted Execution Environment, TEE）基于硬件实现隐私计算，在计算机硬件平台上引入安全芯片架构，构建一个安全的硬件区域，各方数据统一汇聚到该区域内进行计算，通过其安全特性提高终端系统的安全性。可信执行环境目前有英特尔的 SGX 技术、ARM 的 TrustZone 技术，以及 AMD 的安全内存加密，安全加密虚拟化技术等。

1. SGX 技术

SGX (Software Guard Extensions) 通过提供一系列 CPU 指令码，允许用户代码创建具有高访问权限的私有内存区域 (Enclave-飞地)，包括 OS, VMM, BIOS, SMM 均无法私自访问 Enclave，Enclave 中的数据只有在 CPU 计算时，通过 CPU 上的硬件进行解密。同时，Intel 还提供了一套远程认证机制 (Remote Attestation)，通过这套机制，用户可以远程确认跑在 Enclave 中的代码是否符合预期。

2. TrustZone 技术

ARM TrustZone 通过对原有硬件架构进行修改，在处理器层

次引入了两个不同权限的保护域：安全域和普通域，任何时刻处理器仅在其中的一个环境内运行。这两个域之间有硬件隔离和不同权限等属性，为保护应用程序的代码和数据提供了有效的机制。通常正常域用于运行商品操作系统，该操作系统提供了正常执行环境(REE)，安全域则始终使用安全的小内核(TEE-kernel)提供 TEE，机密数据可以在 TEE 中被存储和访问。

3. AMD Zen 技术

AMD 于 2016 年提出硬件内存加密机制 SME 与 SEV 并将其与现有的 AMD-V 虚拟化技术进行了结合。在原有的 CPU 中引入了新的片上芯片系统并将其与原有的内存控制器进行了结合以支持 SEV 和 SME 的实现。

SME 模式下，AMD 安全处理器会在数据写入内存的时候，使用 AES-128 密钥进行加密，而且每次系统重置后都随机生成新的密钥，运行在 CPU 核心上的任何软件都不可能获取。

SEV 模式下，传统令牌加密系统中虚拟机监视器(Hypervisor)泄密的问题将不复存在。监视器和运行在同一机器上的客户端资源、不同客户负载完全隔离，代码和数据将分别标记、独立加密，访问的时候如果加密标记错误，就只能看到加密状态。

国内的海光 x86 CPU、华为 ARM CPU、飞腾 ARM CPU 等都实现了 TEE。蚂蚁推出的 HyperEnclave 是一个统一的 Enclave 平台，目前已经在蚂蚁大规模落地，支撑了多个关键业务，同时，联合合作伙伴，在金融、医疗等领域开展了部署和实施工作。

（四）同态加密

1. 总体介绍

同态加密 (Homomorphic Encryption) 是基于数学难题的计算复杂性理论的密码学技术，能实现密文间的多种计算功能，即先计算后解密可等价于先解密后计算，其结果与用同一方法处理未加密的原始数据得到的输出结果是一致的。1977 年 RSA 算法诞生，但由于当时同态加密算法本身的局限性，在应用上非常受限。直到 2009 年，密码学家们成功构建出了全同态加密方案，这让整个密码学领域重燃在此领域进一步突破的信心。相比于多方安全计算，同态加密在行业上的产品落地上相对较难。但在一些特定应用场景，如机器学习等，通过对算法的适配优化，亦能满足实际业务需求^[7]。

2. 全同态和半同态

根据支持的同态运算类型可以将同态加密方案分成两大类：半同态加密和全同态加密。其中半同态加密是对不具备全同态能力的同态加密方案的统称，即只能支持同态计算一种运算，或者加法，或者乘法。全同态加密指的是能够同时支持加法与乘法运算的同态加密方案。但全同态加密理论难度较大，在密码学界被认为是研究难度最高的研究方向之一。

（1）半同态加密

RSA: 基于大整数分解问题，仅支持乘法同态运算，因其为确定性加密方案，故不满足语义安全性，实际应用较少；

Paillier: 基于 n 次剩余问题，仅支持多次加法同态和数乘运算，目前应用最广。

ElGamal: 基于离散对数难题问题，仅支持乘法同态运算，可使用椭圆曲线加密算法。

(2) 全同态加密

根据对自举 (Bootstrapping) 技术以及计算密钥的使用可以将目前的全同态加密方案分成三代：

第一代同态加密方案通过稀疏子集和假设来压缩解密电路，并应用自举技术实现全同态加密；第二代同态加密方案放弃了计算量过大的自举技术，通过优化方案的代数结构以及一些快速运算算法提升了方案的效率与安全性；第三代同态加密方案无需使用计算密钥的辅助即可进行全同态运算，由于计算密钥的尺寸一般非常大，这一改进极大提升了同态运算的计算效率。鉴于第一代同态加密方案应用较少，本文主要介绍二、三代同态加密方案。

BGV 方案: 第二代全同态方案，现有 Helib 等开源库支持。方案中依次使用模数转换控制噪音的增长，在实际应用中无需启用计算量过大的自举。

CKKS 方案: 第二代全同态方案，支持浮点数计算，非常适合统计和机器学习应用，现有 SEAL 等开源库支持。方案加密时对明文进行舍入，解密时输出满足精度要求的近似值。

GSW 方案：第三代全同态方案，同态加法和同态乘法都只是通过做简单的矩阵加法和乘法来实现，无需计算密钥参与同态运算，现有 TFHE 等开源库支持。

（五）差分隐私

差分隐私（Differential Privacy, DP）在不损害个人隐私的前提下实现了最大限度利用数据资源的核心诉求，通常用于解决单个查询的隐私保护问题。但在实际中，经常需要面临多条隐私计算组合或在同一数据集重复执行相同的隐私计算的情况。

差分隐私分为全局差分隐私和本地化差分隐私，全局差分隐私可以实现加入很小的噪声保护数据集中所有用户的隐私，但是需要所有用户将未经处理的原始数据直接存储在一个可信服务器上。本地化差分隐私允许用户在本地图机化原始数据后再发送给服务器，使得用户得到了更强的隐私保护，但这一过程中加入的噪声远大于全局差分隐私。

差分隐私最初是为数据库中的安全查询设计的，随着联邦学习的兴起，差分隐私也被用于联邦学习算法。目前最常见基于差分隐私的机器学习算法是 differentially private stochastic gradient descent（DPSGD），即基于差分隐私的随机梯度下降法。该算法通过干预模型用来更新权重的梯度来保护训练集的隐私，每次迭代中向梯度添加噪声，在多个批次训练以后，噪声自然会被抵消。

（六）零知识证明

零知识证明 (Zero-Knowledge Proof) 是由 S. Goldwasser, S. Micali 及 C. Rackoff 在 20 世纪 80 年代初提出的。指的是证明者能够在不向验证者提供任何有用的信息的情况下,使验证者相信某个论断是正确的。零知识证明实质上是一种涉及两方或更多方的协议,即两方或更多方完成一项任务所需采取的一系列步骤。证明者向验证者证明并使其相信自己知道或拥有某一消息,但证明过程不能向验证者泄漏任何关于被证明消息的信息。

零知识证明起源于最小泄露证明。设 P 表示掌握某些信息,并希望证实这一事实的实体,设 V 是证明这一事实的实体。假如某个协议向 V 证明 P 的确掌握某些信息,但 V 无法推断出这些信息是什么,我们称 P 实现了最小泄露证明。不仅如此,如果 V 除了知道 P 能够证明某一事实外,不能够得到其他任何知识,我们称 P 实现了零知识证明,相应的协议称作零知识协议。

（七）数据去标识化及脱敏

1. 总体介绍

脱敏技术一般常用于敏感数据的处理,将某些敏感信息通过脱敏规则进行数据的变形,从而降低数据敏感度,减少敏感数据被精确识别的风险,从而实现敏感数据的保护。数据脱敏在保留一定的数据可用性、统计性等基础上,通过失真等变换降低数据

敏感度,脱敏数据需要进行传输通信,而在传输过程中或者之后,攻击者或数据获取方仍可通过特定的技术手段对脱敏后的数据进行推理,进而获取部分乃至全部原始信息。

去标识化属于脱敏技术中的一种,一般是针对个人信息的脱敏处理。通过去标识化计算,使其在不接触额外信息的情况下,达到无法识别个人信息主体的效果。

两种技术所针对的数据范畴虽然不同,但是在实现时所采用的技术方案基本是一致的,计算性能高,适用于大数据量处理,但其可追溯性差,数据脱敏后的去向和使用难以从技术上有效控制,只能作为隐私计算应用过程的辅助手段用来隐藏数据信息。

2. 主要技术方案

主要的技术方案包括以下几大类:密码技术、假名技术、抑制技术、泛化技术、随机化技术、统计技术、数据合成技术。

(1) 密码技术。通过密码学的加密算法将数据进行加密,以完成变形脱敏。采用密码技术脱敏后的数据是可以还原的。当需要还原的时候,用相同的算法并输入密钥,即可完成还原。常用于脱敏的加密算法可以分为确定性加密和随机性加密两类。确定性加密的特点是相同的明文使用相同的密钥加密后都对应到相同的密文,如常用的对称加密算法和非对称加密算法,又如对密文格式有特殊要求的保序加密和保留格式加密;而随机性加密的特点是,相同的明文会每次加密后都产生不同的密文,例如同

态加密。

(2) 假名技术。直接使用假名进行替换来完成脱敏。假名化技术一般采用某种计算规则由原始数据参与计算后生成假名数据，或直接随机生成假名数据。利用假名技术脱敏后的数据无法直接进行还原，但是可以通过建立原始数据-假名数据的映射表来实现假名数据的还原。如果遇到需要将多份假名化处理后的数据进行关联的情况，则这几份数据需要采用相同的计算规则进行假名化，或通过同一张映射表进行假名化，否则就会出现数据无法打通的情况。假名技术最常用的计算规则就是各类散列算法。但是，由于原始数据空间的可举性，攻击者可通过彩虹表的攻击形式对假名化的结果进行反向还原。为应对这种攻击，可在计算散列值时加入 salt，以提高破解难度。

(3) 抑制技术。对需要脱敏的数据项进行删除或进行屏蔽。抑制技术适合用于具有可识别性特征的属性字段的脱敏处理。这类字段虽然不具有唯一可识别性，但是结合其他信息就能够具体识别到某一特定的信息主体。抑制技术可以采取全部抑制，即删除（清空）处理，也可以采用部分抑制技术，对字段中的部分信息进行屏蔽或遮掩。采用抑制技术脱敏后的数据是无法还原的。

(4) 泛化技术。通过降低数据集中所选属性粒度来实现脱敏。泛化技术的目标是减少属性的唯一值，尽可能的消除原始数据的唯一性，使原始数据中的多个数值都对应到泛化后的同一个值上。常用的泛化技术有分层、取整等，采用泛化技术脱敏后的

数据是无法还原的。

(5) 随机化技术。通过随机化修改属性的数值以达到脱敏效果。对于随机化后的数据集里的单条数据，其属性特征已发生了改变，因此很难结合其他数据属性推断出特定的信息主体。常用的随机化技术包括噪声添加、置换等。置换技术在不改变数值的前提下，将数据集里所选属性的值进行重新排列，保持了数据集的统计特性。在进行噪声添加时往往也会尽量保证数据集的统计特性。采用随机化技术脱敏后的数据是无法还原的。

(6) 统计技术。利用统计学的方法，将属性数据进行处理来实现脱敏，同时又能够保留该数据集的统计学特性。常用的统计技术有数据抽样和数据聚合。数据抽样是在数据集中选择具有代表性的子集来对原始数据集进行分析和评估。数据聚合则是利用统计值替代属性的具体数值，这样就无法反映出单条数据记录的特征，但数据集整体的统计特性没有发生改变。采用统计技术脱敏后的数据是无法还原的。

(7) 数据合成技术。以人工的方式生成数据，使其符合该属性的取值范围。数据合成的技术常用于测试数据的生成。往往测试环境中需要一些数据进行系统测试和验证，而真实数据用于测试环境可能会造成敏感数据泄漏。在此种场景下数据合成技术可帮助解决此问题。

（八）相关技术结合

1. 隐私计算与区块链

为了保证数据的隐私性，隐私计算实现了在多方协作计算过程中对于输入数据的隐私保护，但是原始数据、计算过程和结果均缺乏可验证性。而区块链是一种由多方维护，使用密码学技术保证数据传输安全、数据一致性和防篡改性的分布式账本技术。虽然多方安全计算符合区块链去中心化的特性以及解决信任问题的终极目标，但双方的实践方式有所不同。两者难以做到有机结合，只能通过技术应用提升自身某一特性。目前，隐私计算与区块链的结合上具体分为两种。

（1）区块链技术在隐私计算中的应用

隐私计算可以借助区块链技术实现操作记录存证，通过区块链的可追溯性获得记录可验证的特性。以信贷风控为例，隐私计算可以实现多企业、机构之间的高效安全数据共享。各方建设数据平台时，无需再将来自其他机构的数据全部迁移至自身平台，而只需将运算模型或规则布署在各机构的数据域内，根据业务请求实时进行加密计算，实时调用。在整个计算过程中涉及的数据摘要、判断结果均可加密存证于区块链，便于后续的审计、监督。

区块链可应用到隐私计算过程中数据生命周期的各个环节中，实现全程闭环的安全和隐私服务，操作和处理记录可上链保存、不可被篡改。区块链可以解决数据共享参与者身份及数据可

信问题，一定程度避免主观作恶、合谋推导、数据造假等。区块链可以提升数据共享流通协作效率，降低隐私计算应用成本。比如数据持有者可将共享数据目录、数据使用申请、数据使用审批、数据使用审计等功能上链。

（2）区块链中的隐私计算

区块链可以通过采用隐私计算技术来提升自身的数据保密的能力，以适应更多的应用场景，通过同态加密等技术，在保证数据隐私性的同时，完成区块链共识算法同步或智能合约执行。比如对区块链上的账户金额数值进行加法同态加密，验证节点在不需知道交易金额的具体数值或者其他任何隐藏信息的情况下，可以对密文进行正确的加法操作，这样区块链上所有的用户余额和交易金额都以同态密文的形式存在。除了拥有私钥的可信第三方机构外，所有节点都只能验证交易而无法得知具体数值，这将有效的保护用户的账户隐私。还可以将区块链上链过程中的私钥签名环节放入 TEE 执行，通过 TEE 提供的可信环境，保证了用户私钥没有泄露，签名过程没有被篡改，从而为整个区块链网络提供了安全性保障。

2. 隐私计算与深度学习

隐私深度学习技术可以看做是广义的隐私计算技术在深度学习技术领域的特定应用，核心目标是实现在不泄露多参与方私有数据隐私前提下实现深度学习模型的训练和推理，同时确保

“可用”和“不可见”两个目标。

传统的深度学习技术以数据归集后形成的大数据资源池为基础，往往面临因“数据孤岛”问题导致深度学习技术的效能和应用发展受到桎梏。业界主要隐私深度学习技术方向包括基于 MPC 的深度学习路线和联邦学习路线。基于 MPC 的深度学习是通过可证明安全的密码算法、密码协议实现交互数据和计算结果的隐私保护。联邦学习是一种分布式机器学习架构，通过数据切分和模型迁移实现数据不共享前提下协作构建机器学习模型，该方法的安全性目前尚难严格证明。因此，基于 MPC 的深度学习路线是学术界和产业界发展的重要方向。

运用基于 MPC 的深度学习，一般基于算术电路，采用秘密分享技术，对数据进行碎片化处理，设计各类基础算子和衍生算子用于构建复杂的深度学习模型。基于算术电路和秘密分享的 MPC 协议，适用于机器学习类复杂计算任务。基于 MPC 的深度学习框架如图 2-7 所示。

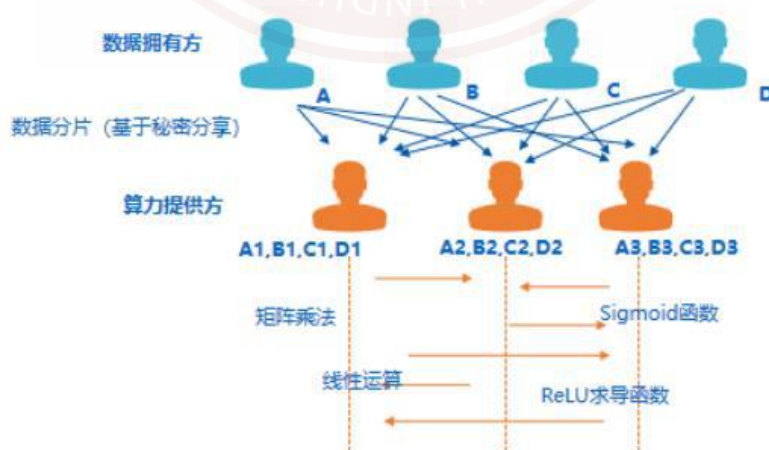


图 2-7 基于 MPC 的深度学习框架

3. 隐私计算与知识图谱

知识图谱是由数字、文字、图像、符号等，经过筛选、分析、归纳、总结等知识组合构建形成，提炼出各种有用的知识，进而可以构建机器的先验知识，用于智能搜索、深度问答、智能决策等场景任务推理。

目前知识图谱的应用都是基于单一完整图谱的理想状态进行设计，但在实际应用场景中，构成知识图谱的知识往往散落在不同的机构或个人手中，形成一个个数据孤岛。如果仅基于自身数据构建来构建图谱，由于数据量原因，推理准确率往往较低。同时，出于自身利益和数据合规性的考虑，各机构难以直接通过数据共享的形式，将数据集中起来形成完整的知识图谱，进而进行推理。通过知识的表示学习，在不泄漏彼此数据的前提下，对多个参与方知识进行抽象，并完成联邦融合，最终完成协同知识联邦推理。

(1) 知识推理技术方案

(a) 知识抽象表示，各个机构组织，在日常运营过程中，产生各种各样的知识，组合构建为知识图谱。通过图神经网络算法，可以把图节点特征及相关结构信息，抽象表示为低维向量。见图 2-8 左半部分；

(b) 知识联邦融合，各个参与方依据自身知识，经过知识编码，形成对相同实体的各自低维向量表示。第三方接收各参与方抽象知识，进行联邦融合，得到综合后的向量表示。第三方获

取的抽象知识不包含敏感数据，也不可反推出各个参与方的图谱原貌。见图 2-8 右半部分。

(c) 知识联邦推理：根据任务不同，可以选择不同的推理方式。对于节点分类任务，得到综合后的向量表示，可直接激活函数得到最终的节点类别。

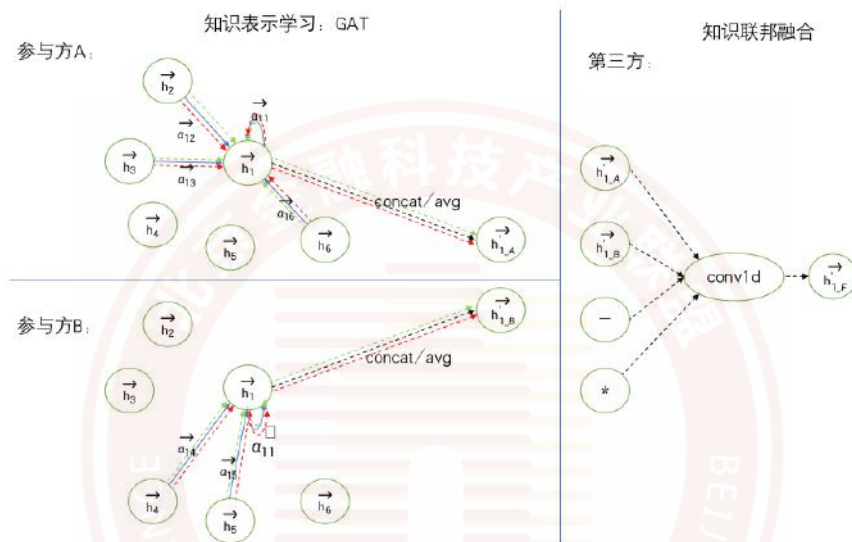


图 2-8 基于知识联邦的知识推理方法

(2) 本技术的优势

(a) 把各参与方连接起来形成知识联邦，协同进行知识推理；

(b) 弥补现有隐私计算技术对于非结构化数据的处理能力的不足；

(c) 在训练与推理中，只传输抽象后的知识表示向量，保证各参与方数据安全及隐私；

(d) 经实验验证，试用知识图谱推理方法协同分类准确率

远超单节点准确率，接近集中化分类准确率。

金融行业的风控应用中，为了提升风控效率，降低人力决策效率低等信贷审核成本，除了利用传统风控经验规则、基于机器学习的风控模型，图数据同样也有很多技术应用。以反洗钱业务为例，机器学习基于关系型数据建模无法挖掘出客户之间的关联信息和推导资金轨迹关系。而关联账户之间是否存在交易链路的闭环，是反洗钱的重要风险挖掘的方法之一。图学习可以显示已知关系，揭示隐藏的联系、网络和集群，可以有效地建立资金流动轨迹和参与洗钱活动的实体-关系网络，但对于金融企业而言，有限的关联数据会导致关联关系往往在关键的节点断开，从而无法最大化图的知识反馈，很多闭环信息也就无从挖掘，效果难以满足要求。

通过提升本图数据的节点边关系结构，并且不泄露互相之间的用户数据隐私，金融机构能够分配更多的时间和资源来审查更高的风险警报，大幅降低了反洗钱的挖掘和节省时间，降低成本，并且通过快速实时的风险提示，在确保审核质量的前提下，降低遗漏高风险行为的风险。

4. 隐私计算与自然语言处理

自然语言处理是研究人与计算机交互中的语义理解问题的技术，解决的是人机交互中“听得懂”的问题。其技术应用流程包括语料获取及预处理、特征提取、模型训练、模型评估与应用。

自然语言处理算法高度依赖于用户的本地数据，例如文本信息、文档及其标签、问题和选择的答案等，这些数据既可能位于个人设备上，也可能位于不同机构更大的数据仓库中。在真实的场景中，用户的自然语言数据是敏感的，可能包含隐私内容，很难训练出一个健壮模型来造福用户。将联邦学习应用到自然语言处理领域中，有助于开发一些隐私保护、个性化的语言模型。

最经典的自然语言处理结合隐私计算的例子是谷歌的利用移动设备用户数据进行的横向联邦学习，基于移动设备用户频繁键入的单词来学习词库外（Out-of-Vocabulary, OOV）单词。词库外单词是指不包含在用户移动设备的词库表中的词汇。词库表中缺少的单词无法通过键盘提示、自动更正或手势输入来预测。从单个用户的移动设备学习 OOV 单词来生成模型是不切实际的，因为每个用户的设备通常只会存储有限大小的词库表。收集所有用户的数据来训练 OOV 单词生成模型也是不可行的，但 OOV 单词通常包含用户的敏感内容。因此 2019 年，谷歌实现了首个产品级的联邦学习系统，主要侧重针对 C 端，在移动手机上运行的联邦平均算法和分析。联邦学习可以根据所有移动用户的数据，训练一个共享的 OOV 生成模型，并且不需要将敏感内容传输到中心服务器或云服务器上。

联邦 OOV 模型训练流程如图 2-9 所示。首先每一台移动设备从服务器下载共享模型，接着每一台移动设备基于用户输入的内容，训练共享模型，再通过安全协议上传至服务器。最后服务器

从移动设备收集更新，聚合这些更新并以此改良共享模型。迭代执行这个步骤来训练共享 00V 生成模型，直到模型收敛为止。

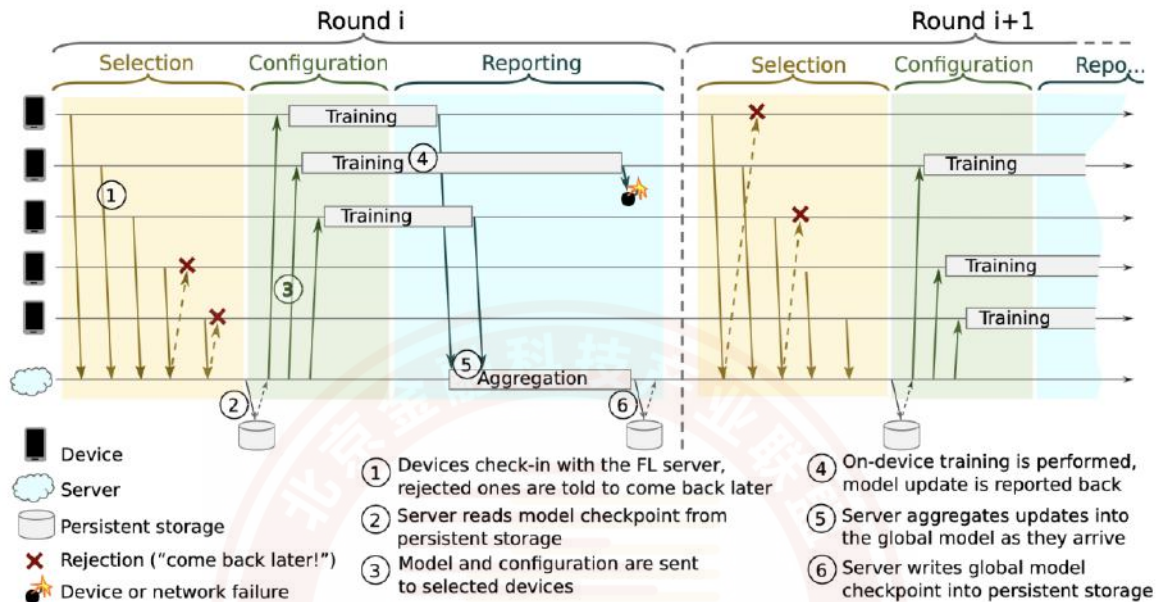


图 2-9 联邦 00V 模型训练流程

隐私计算还可以与 NLP 中各种流行的文本分类、序列标记、对话系统、seq2seq 生成和语言建模等任务结合。例如，目前落地最多的有：基于 FL 的键盘下一字预测；使用 Text-CNN 的句子级文本意图分类；使用来自多方的医疗数据对 Bert 进行预训练和微调，而无需将所有数据聚合到同一位置；将与联邦学习结合的方法来训练高质量的语言模型，这些模型可以优于在没有联邦学习的情况下训练的模型；结合联邦学习在医学上完成关系提取和医学名称实体识别方面等。

未来，自然语言处理结合隐私计算还有一些挑战性的工作：丰富基于隐私计算的新型应用和模型，如医疗领域、翻译领域等；

设计优化 NLP 中 non-IID 问题，有效利用未标注数据；提高大型语言模型的通信性能；提高现有的隐私保护的能力，如在 NLP 领域防御后门攻击、中毒攻击等。

5. 隐私计算与机器视觉

深度神经网络模型是现代机器视觉的主流技术。支持深度神经网络模型的一个重要的因素是海量的高质量标注数据。而获得这些数据往往是成本高昂的。因为对数据的筛选和标注需要大量的人力和物力成本。在专业性高的领域（比如金融，医疗等领域）尤其如此。解决高质量标注数据难以获得的一个比较普遍的方案是数据共享。然而，由于用户隐私、监管风险、缺乏诱因等原因，许多企业并不愿与其他企业直接共享数据。

联邦学习是替代数据共享，实现知识共享的一种新的学习范式。在联邦学习中各个参与方（移动设备，企业或机构）的自有数据不出本地，通过在加密机制下的参数（模型参数或计算中间结果）交换方式，各个参与方能够在不违反数据隐私的法律法规情况下，协同地建立一个联邦模型。目前，联邦学习已经在机器视觉领域中的多个场景落地。

虽然联邦学习能够使所有参与方在不公开己方数据的情况下，协同地训练一个联邦模型。但是，要得到能够满足所有参与方需求的联邦模型，通常需要参与方的数据是服从独立同分布的。然而，在很多实际场景中，各个参与方所拥有的数据之间存

在着不小的差异，比如，各参与方的数据分布情况可能有较大差异，各参与方拥有的数据量级可能差距很大，或者各参与方拥有不同模态的数据等。

联邦学习可以结合各种迁移学习技术来消除或者减弱参与方之间的数据异构性，使联邦学习可以应用于更广的业务范围。将联邦学习与迁移学习技术的结合统称为联邦迁移学习。联邦迁移学习已经广泛应用于联邦学习的各个场景中。

在机器视觉领域，联邦迁移学习被应用于医学图像分析场景。受限于患者数量或者病理类型，单个医疗机构（包括医院）通常只有少量的有标注的医学图像，同时各个医疗机构收集得到的医学图像在特征分布上也存在差异。因此，依靠传统的联邦平均算法将得不到满足需求的联邦模型。有研究人员提出一种差异感知的联邦学习方法来削弱不同数据源之间的医学图像分布差异，以帮助目标领域的参与方建立精确度高的前列腺癌分类模型。其核心思想是通过基于隐私保护的生成对抗模型将各个参与方的原始数据映射到目标领域的图像空间，从而最小化各个参与方之间的数据差异。然后，在这个公共的图像空间中进行联邦学习，以得到在目标领域具有高精确度的模型，并通过应用混合专家模型和对抗领域对齐技术来解决联邦功能性磁共振成像分析中的数据异构问题。

（九）技术对比分析

1. 技术安全性及验证

从安全性、计算准确性、计算性能以及通用性（应用场景和领域）等角度进行比较，各隐私计算技术各有利弊：

差分隐私技术通过增加噪声来保护数据隐私，计算性能也很高，但噪声带来的偏差使得结果准确性降低。

可信执行环境为程序、数据提供了一个安全可靠的环境，其性能及通用性具有较大优势，但信任链绑定 CPU 厂商，且理论上存在侧信道攻击的可能性。

联邦学习通过数据不出本地、只交换加密中间模型参数的方式实现多方安全建模。联邦学习技术的安全性基于相信无法通过中间模型参数推断出原始数据，但此结论没有密码学保证，因此联邦学习需要和其他密码技术结合来保证安全性。

与上述技术相比，多方安全计算基于密码学安全，其安全性有严格密码理论证明，不以信任任何参与方、操作人员、系统、硬件或软件为基础，同时计算准确度高，并支持可编程通用计算。

2. 性能评估

不同的隐私计算技术，具备不同的性能特点，适用于不同场景。具体论述如下：

基于密码学的多方安全计算及同态加密等方法更适用于数

据量适中但保密性要求较高的重要数据应用。从技术上看，计算性能问题是应用的一大障碍。随着应用规模扩大，采用合适的计算方案保证运算时延与参与方数量呈现线性变化是目前各技术厂商面临的一大挑战。

联邦学习适用于数据量大的模型训练。联邦学习由于其具有分布式训练和联合训练的特点，一方面能够解决训练阶段数据特征单一的问题，从而获得一个性能更好的、优于利用自己本身数据集所训练出的模型。另一方面，各参与方只需在本地利用各自数据集进行训练，数据体量未增加，算力成本压力较小。但模型的梯度及权重等信息涉及到加密及传输，会导致相比于明文融合数据的计算来说效率降低。

差分隐私通过增加噪声减少计算结果对隐私的泄露，一般与其他技术结合使用。在多方安全计算中，采用可计算的差分隐私能大大降低多方安全计算的计算复杂度和通讯量。但就目前多个研究，差分隐私由于其技术机制原理，算法输出结果并不精确，学术界还在进行更高效的差分隐私技术研究。因其原理为添加噪声，相当于在明文做计算，计算效率高。

可信执行环境技术具有通用和高效的特点，不仅可以无缝支持通用计算框架和应用，而且计算性能基本可匹敌明文计算（比如正常的 Linux 计算应用）。因此，可信执行环境技术由于性能优势而更适用于复杂、数据量大的通用场景和通用算法，如大数据协作、人工智能框架数据保护、关键基础设施保护等，但是目

前的安全性受限于硬件的设计与实现。

可证去标识同样适用于数据量大、实时性要求高、数据出域的应用场景。这种技术确保数据去标识后，数据接收方无法重新识别或者关联个人信息主体。可证去标识对参与计算的多方数据可信去标识管控，确保所有计算基于去标识化数据展开；在结果输出阶段对输出数据进行原始数据拥有主体及用户的双重确权，实现了价值输出时各方权益可保障。

3. 技术差异对比

常用的隐私计算技术的差异对比如表 2-2 所示。

表2-2 隐私计算技术对比^[8]

技术	性能	通用性	安全性	可信方	整体描述
多方安全计算 (MPC)	低~中	高	高	不需要	通用性高、计算和通信开销大、安全性高，研究时间长，久经考验，性能不断提升
可信执行环境 (TEE)	高	高	中~高	需要	通用性高，性能强，开发和部署难度大，需要信任硬件厂商
联邦学习 (FL)	中	中	中	均可	综合运用 MPC、DP、HE 方法，主要用于 AI 模型训练和预测
同态加密 (HE)	低	中	高	不需要	计算开销大，通信开销小，安全性高，可用于联邦学习安全聚合、构造 MPC 协议
零知识证明 (ZKP)	低	低	高	不需要	广泛应用于各类安全协议设计，是各类认证协议的基础
差分隐私 (DP)	高	低	中	不需要	计算和通信性能与直接明文计算几乎无区别，安全性损失依赖于噪声大小

经典多方计算不需要可信第三方，通用性高。同态加密基于密码学理论，通用性强便于理解，但由于其底层的复杂计算因此效率较低。可信执行环境由于计算依赖硬件，因此计算性能高通用性好，但需要对硬件厂商高度信任。混淆电路由于多轮多方的信息交互导致通讯效率较低。差分隐私基于概率统计学，主要用于统计，计算性能高，应用在较特定的场景。零知识证明多与区块链结合使用，主要用于链下证明，链上认证，由于底层计算复杂，因此效率不高。联邦学习和共享学习比较相似，均是结合隐私计算与机器学习的综合应用，但由于其底层运用技术的不同做了区分，联邦学习偏重使用基于密码学的隐私计算技术，如同态加密、多方安全计算等；基于 TEE 的共享学习偏重使用基于硬件的隐私计算技术，如 TEE、硬件加速等。

总体来说，不同的隐私计算技术在计算性能、通用性、安全性、安全机制、可信方等方面都存在差异，在实际选择时需要综合考虑行业特点、业务场景、客户需求等，从而扬长避短，充分发挥隐私计算的价值。

三、隐私计算金融应用背景

随着金融信息化的不断发展，金融行业积累了大量高质量的敏感数据，事关国家经济命脉和公民合法权益。因此，金融数据的安全与风险防范一直是金融监管部门关注的重点，国家也相继出台金融数据安全相关政策与法律法规，对金融数据安全的重视程度与日俱增。为实现业务能够更好地开展，金融行业产生了跨机构、跨行业进行数据协作的需求。隐私计算技术能够在保证数据安全和信息隐私的前提下，联合多方数据进行分析挖掘，满足多样化场景下更高的数据要求，得到重点关注，但目前仍处于一个较早期的阶段，在金融行业应用中仍面临一些风险和挑战，需要在不断的实践中成熟。

（一）政策与法律法规

2020 年 4 月，国务院印发《关于构建更加完善的要素市场化配置体制机制的意见》（中发〔2020〕9 号），明确把数据列为生产要素，并要求“加强数据资源整合和安全保护”“制定数据隐私保护制度和安全审查制度”。

2020 年 5 月，国务院印发《关于新时代加快完善社会主义市场经济体制的意见》（中发〔2020〕10 号）中明确提出：“加快培育发展数据要素市场”“加强数据有序共享，依法保护个人信息”。

2020 年 12 月，国家发改委联合 3 部委发布《关于加快构建全国一体化大数据中心协同创新体系的指导意见》（发改高技〔2020〕1922 号），以深化数据要素市场化配置改革为核心，优化数据中心建设布局，推动算力、算法、数据、应用资源集约化和服务化创新。

2019 年 8 月，中国人民银行发布《金融科技（FinTech）发展规划（2019-2021 年）》（银发〔2019〕209 号文印发），提出“科学规划运用大数据”“稳步应用人工智能”“增强金融业务风险防范能力”“加大金融信息保护力度”。

2021 年 12 月，中国人民银行发布《金融科技发展规划（2022—2025 年）》（银发〔2021〕335 号文印发），明确提出从强化数据能力建设、推动数据有序共享、深化数据综合应用、做好数据安全保护方面充分释放数据要素潜能。

目前金融业逐渐向“智能科技+金融”数字化转型，以银行为主的传统金融机构在深化转型，全面发挥自身产品、网点、客户基础等优势，在对产品进行创新的同时，积极运用大数据和人工智能技术全面提升业务智能化能力。但在这个转型进程中，数据监管越来越严格。欧盟保护公民隐私的《通用数据保护条例》（General Data Protection Regulation，简称 GDPR）于 2018 年生效，美国《加州消费者隐私保护法案》（California Consumer Protection Act，简称 CCPA）于 2020 年生效，中国《中华人民共和国网络安全法》于 2017 年发布，《中华人民共和国数据安全

本法》《中华人民共和国个人信息保护法》于 2021 年发布。法律的发布对用户隐私保护起到了重要的作用，这些法律法规中对数据和信息的保护条例对大数据、人工智能应用在各个场景中的数据处理模式提出新的挑战，不能处理好数据服务和用户隐私保护之间的关系，将极大阻碍信息化的发展，甚至给企业和社会带来不良影响，因此“隐私安全保护”是一个智能金融进一步发展面临的必须要解决的问题。但同时法律法规为智能金融指出了一条可发展之路：鼓励企业利用技术手段，解决数据隐私保护问题，促进安全的数据流通。

另外“告知-同意”是法律确立的个人信息保护核心规则^[9]。个人信息处理者在取得个人同意的情形下方可处理个人信息，个人信息处理的重要事项发生变更，应当重新向个人告知并取得同意。

《个人信息保护法》规定了个人信息处理的基本原则：第一，合法正当诚信原则；第二，处理必要原则；第三，目的特定原则；第四，知情同意原则；第五，个体参与原则；第六，质量保证原则；第七，公开透明原则；第八，安全保障原则。

隐私计算的目标主要是防止数据中隐私信息的泄露，因此隐私保护技术最主要的性能指标为隐私保护强度。隐私计算的技术目标与《个人信息保护法》的立法目的相符，而隐私计算技术对个人信息数据的保护路径与《个人信息保护法》的主要原则也存在高度结合。隐私计算提供了多种信息保护方法，数据安全性高，

同样需要遵守个保法及相关配套法律法规，妥善处理不同应用场景所涉个人信息处理及隐私计算应用环节的各方权利义务及法律责任，避免使用隐私计算从事违法违规活动^[10]。

根据《网络安全法》及《民法典》相关规定，数据处理者在处理数据时应公开收集、使用规则，并经用户同意。当前，各应用网站隐私政策中明确告知了对用户个人信息的收集、使用、管理等方面的处置办法，其中使用包括但不限于共享、转让和披露等。从隐私计算的特点来看，数据合作方通过隐私计算技术实现数据分析与建模，不需实际流转数据，且处理过程中的数据都进行了匿名化处理，或不需要获得用户授权同意。但在数据采集阶段，数据合作各方仍需获得用户授权同意。此外，个人信息的匿名化标准尚存争议，因此仍需做好告知同意的授权管理。隐私计算在本地服务器中建模的行为也存在用户授权的问题。即使企业在采集数据时通过隐私政策取得了用户对本地建模行为的授权，但该授权仍需保持在与数据实际处理目的直接或合理关联的范围内^[11]。

（二）国内外相关标准

隐私计算技术基于密码学、统计学等基础学科理论并利用工程优化，有效实现了数据“可用不可见”的创新模式，从技术手段上打破了“数据孤岛”的合作壁垒。随着隐私计算技术近几年的快速发展和落地应用，国际与国内隐私计算的相关标准相继制

定发布。

在国际上，电气和电子工程师协会（IEEE）标准组于 2020 年 3 月发布了《联邦学习系统架构和应用指南》（IEEE 3652.1-2020 - IEEE Guide for Architectural Framework and Application of Federated Machine Learning），该标准对联邦学习的定义、概念、分类、算法框架规范、使用模式和使用规范等方面都进行了系统性阐述，并对联邦学习在 To B（企业端）、To C（用户端）以及 To G（政府端）不同情境下的场景分类，建立了联邦学习的需求分析模板，厘定了联邦学习性能及安全测评准则。此外 IEEE 标准组《多方安全计算推荐实践》（IEEE 2842-2021 - IEEE Recommended Practice for Secure Multi-party Computation），国际标准化组织 ISO《多方安全计算标准》，国际电信联盟 ITU《多方安全计算技术指南》等均在制定当中。

在国内，中国人民银行于 2021 年 3 月发布《金融业数据能力建设指引》（JR/T 0218-2021）标准，明确了金融业数据工作的基本原则，从数据战略、数据治理、数据架构、数据规范、数据保护、数据质量、数据应用、数据生存周期管理等方面划分了 8 个能力域和 29 个对应能力项，提出了每个能力项的建设目标和思路，为金融机构开展金融数据工作提供全面指导。该项政策将进一步促进隐私保护技术在金融领域的应用和推广^[12]。中国通信标准协会分别在 2020 年、2021 年发布以信通院牵头制定的四

个关于数据应用与安全的团体标准：《基于多方安全计算的数据流通产品 技术要求与测试方法》（T/CCSA 72-2020）《基于联邦学习的数据流通产品 技术要求与测试方法》（T/CCSA 0503T-YD-2020）《基于可信执行环境的数据计算平台 技术要求与测试方法》（T/CCSA 47-2020）《区块链辅助的隐私计算技术工具 技术要求与测试方法》（T/CCSA 04-2021），从功能、性能、安全性方面对隐私计算的几个主要技术路线进行规范。中国人民银行于2020年11月24日正式发布了《多方安全计算金融应用技术规范》（JR/T 0196-2020），其中规定了多方安全计算技术金融应用的基础要求、安全要求、性能要求等，该标准适用于金融机构开展相关产品设计、软件开发、技术应用等。此外，在2020年底，由支付清算协会联合金融机构、科技公司基于《多方安全计算金融应用技术规范》制定了《多方安全计算金融应用评估规范》（T/PCAC 0009-2021），为多方安全计算产品应用于金融场景提供测评要求与方法。北京金融科技产业联盟在2021年正在组织制定联邦学习金融应用技术规范，从联邦学习的系统技术框架、功能要求、非功能要求、安全要求方面对联邦学习在金融场景的应用进行技术层面的规范，并从认证、管控、计算三个层面对联邦学习系统的互联互通制定框架性要求。

四、隐私计算金融应用场景

金融行业各机构积累了高质量、高价值数据，但数据主要集中于金融领域，且受限于本机构业务范围。从需求角度来看，金融业务需要对客户情况、业务等形成全方位视图，多维度分析才能产生更大的价值。因此，金融机构对同业机构、其他行业机构有着迫切的联合计算意愿。隐私计算技术促进了更加开放的数据价值融合，充分挖掘了数据的潜在价值，既有效支撑了金融业务发展，更好地服务普惠金融、乡村振兴等国家政策，又为金融机构在运营管理、风险控制等方面提供了强有力的支撑。

针对不同的应用场景需求，整体方案采用隐私计算中的一种技术或多种技术组合进行解决。从上一章可以看到，隐私计算所涵盖的联邦学习、多方安全计算、可信执行环境、同态加密、差分隐私等多种技术，在安全、效率、效果等方面有所不同，应用上也各有侧重。本章将从金融行业的应用场景角度，具体阐述隐私计算在智能风控、智能营销、智能运营、隐私信息检索、供应链金融、反洗钱等方向的应用案例。

（一）智能风控场景

银行等金融机构在对客户进行信贷风险控制时，为了得到更加精准的信用风险预测结果，利用外部数据一般采用两种方式：一种是直接购买外部信用分或欺诈分，此种方式受外部数据质量

的限制，无法进一步提升效果；另一种方式为结合客户基本信息、资金流水等自有数据和外部数据联合建模，对其信用风险进行综合评估分析，数据合作涉及到多方客户数据的共享和使用，传统的技术保密方式主要依赖于数据脱敏，但数据泄露事件时有发生，未能很好地保障数据的安全。

银行拥有大量有信贷需求的用户，外部机构有海量的其他类型用户行为数据和场景数据。通过隐私计算，银行无需交换明细级原始数据即可联合其他数据源共同建立风控模型^[13]。

（1）金融机构同政务数据合作风控

案例一：贷后企业风险监测

主要应用技术：联邦学习

根据北京市大数据行动计划，为了发挥大数据在提升城市精细化管理、推动治理能力现代化等方面的重要支撑作用，北京市各委办局的政务信息已陆续汇聚在金融公共数据专区，涵盖工商、医保、社保、公积金等高价值数据，由北京金控集团的全资子公司北京金融大数据有限公司运营并统一对外提供服务。

工商银行与金控集团利用联邦学习技术，在双方数据互不出库、满足用户信息隐私保护以及数据安全的基础上，在企业贷款贷后监测场景中借助金融公共数据专区的企业不动产信息完成联合建模，打造抵押贷款客户不动产信息风险监测产品，为企业贷款业务的动态监测提供数据支持，实现企业抵押贷款贷后监测

预警能力的进一步提升。

产品以企业名称和统一社会信用代码为匹配 ID，联合分行的企业贷款数据（客户贷款余额、期限、状态等）与金融公共数据专区汇聚的企业不动产数据（不动产信息、抵押情况、权利份额等），通过联邦学习平台在双方数据不可见的前提下，共同建立企业贷后监测场景风险评分卡模型，并根据该模型输出的评分转化结果将企业风险状态分为正常、关注、预警三个级别，代表企业的不同资产负债水平和风险级别，以此对抵押类贷款客户在不动产领域进行风险监测及评估。

产品上线后，在贷后管理环节除了可关注企业的经营状况、财务数据以及贷款关联的抵押标的物以外，能够更加全面准确的了解企业当下的不动产信息及抵押情况，掌握其实时资产负债水平，对于临近资不抵债状态的企业及时提示预警，在加强实体经济扶持的同时进一步确保贷款质量。

产品首批对百户存量企业进行监测，对于存在的个别关注企业已安排现场检查沟通，达到有效控制贷款业务风险的目的。工商银行北京分行也是金控集团诸多合作伙伴中，首批唯一一家运用联邦学习技术上线此类产品的单位。后续，分行将继续深化政务数据合作，推动经营管理和业务发展。

案例二：小微贷款类产品风控

主要应用技术：联邦学习

随着数据挖掘应用深入，金融机构开始引入外部数据拓展和

提升数据挖掘的应用范围及使用效果。一方面银企合作升温、外部数据需求增加，另一方面，数据隐私保护已成为很多行业的发展趋势，我国也正推动立法规范数据的留存、使用和流通。在此背景下，民生银行主动拥抱监管要求，在“外部数据价值挖掘”和“数据隐私保护”二者的平衡中，借助联邦学习技术，与其他机构开展小微信贷风控领域的联合建模探索。

本方案如图 4-1 所示，主要包含数据准备、联合建模和解决方案三部分。数据准备阶段，应用隐私求交技术，民生银行与合作企业分别取出各自用户相同但是特征不完全相同的数据，为后续建模做准备；联合建模阶段，选用纵向联邦学习的 XGB 算法进行模型的训练，通过特征维度的增加助力小微风控模型泛化能力的提升；通过不断的调优迭代，最终形成了民生银行小微贷款类产品的风控解决方案。

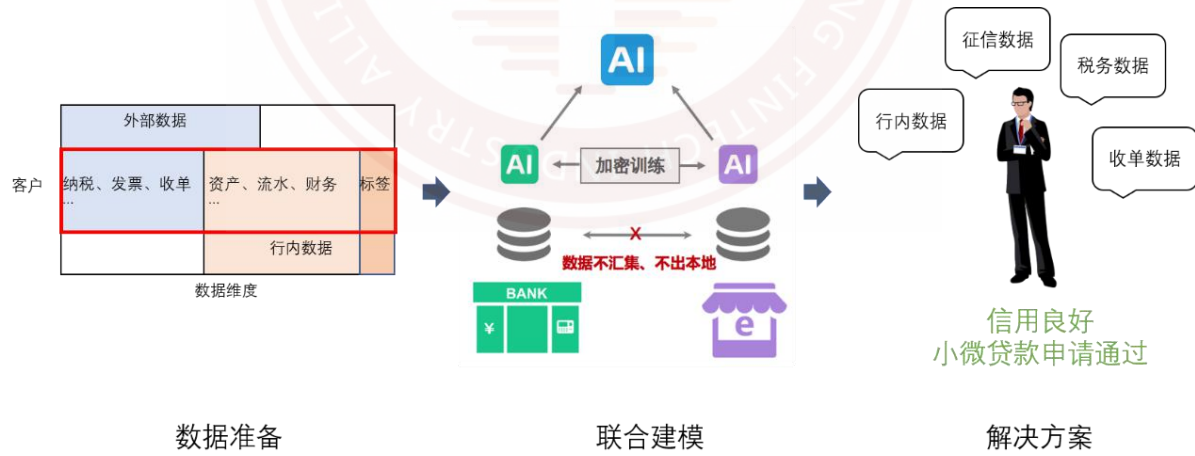


图4-1 解决方案示意

新冠疫情爆发以来，小微企业面临了严峻的经营甚至生存挑

战，民生银行通过应用前沿的联邦学习技术，在保护数据隐私的前提下，探索针对小微贷款类产品的风控解决方案，更高效地为小微企业纾困，助力我国疫情防控稳步向好。

（2）金融机构间数据合作风控

案例一：小微企业新户贷后违约预测

主要应用技术：联邦学习

中国银联是中国银行卡联合组织，通过银联跨行交易清算系统，实现商业银行系统间的互联互通和资源共享，保证银行卡跨行、跨地区和跨境的使用。中国银联已与境内外数百家机构展开广泛合作，银联网络遍布中国城乡，并已延伸至亚洲、欧洲、美洲、大洋洲、非洲等 160 个国家和地区^[14]。

招商银行为快速确认中小微企业新户的基本情况，对中小微企业风险进行高可靠性动态监测，使用银联数据完善企业主模型，用于丰富中小微企业用户画像。通过利用隐私计算技术，评估客户消费能力、信用能力等信息，解决机构无法准确全面了解客户真实能力的痛点，在数据不出本地的监管要求下，双方构建联邦模型，实时匹配。整个建模过程中，机构交互中间参数且不会泄露原始信息，最终构建一个全量模型，助力数据增值，保证数据安全，实现数据真正的流通。模型主要应用在信贷风控、精准营销等场景，帮助机构提升客户服务水平。

招行在行内已有风控模型入模特征基础上进行筛选，同时增

加银联 C 端和 B 端的个人行为数据进行补充构建离线联邦模型，主要沉淀企业主能力模型，用于映射到企业的风控场景，和存量企业客户违约预测。

该案例利用银联全国性数据，扩展银行风控模型入模特征，提升银行现有风控模型效果；利用银联的数据优势，降低企业的获客成本及客户经营风险，提升经济效益，直接、安全并具有可控性。

案例二：小微企业贷前信审

主要应用技术：联邦学习

通过联邦数据网络进行信贷风控增强，从风险源头切入，帮助信贷公司过滤信贷黑名单或明显没有转化的贷款客户，进一步降低贷款审批流程后期的信审成本^[15]。

为响应中央“六稳、六保”，支持小微企业信贷，微众银行以普惠为目标，希望服务更多缺少传统金融服务的小微企业，但因小微企业信贷评审数据稀缺、不全面、历史信息沉淀不足等问题，小微企业信贷业务难实现大规模增长，因此微众银行与中国银联合作，基于联邦学习框架，在双方数据均不出库的情况下使用中国银联交易流数据进行联合建模。联邦形式扩充了数据维度，为信贷决策提供了更有力的支持，促进了业务增长。整体结果表明：采用联邦特征工程对银联数据进行特征筛选，模型效果更好，更稳定，更具解释性。

案例三：排查多头借贷、超额借贷

主要应用技术：多方安全计算-联合统计

某小额贷款公司向个人提供小额贷款服务。当客户向小额贷款公司提出贷款申请时，小额贷公司需要评估借贷风险，排查用户多头借贷及超额借贷的风险。小额贷公司将联合市场上多家有同类服务的借贷公司对用户已借贷的公司数量、贷款总额进行调查。由于客户信息对于贷款公司是宝贵资源，贷款公司需要保护客户隐私不被泄露。

因此基于隐私安全考虑，小额贷公司联合多家同行在腾讯云隐私安全计算平台上，基于安全多方计算技术，在贷前对用户各借贷公司的贷款总额进行联合安全统计。小额贷公司收到联合统计结果后，决定是否向用户发放贷款。

联合安全统计的原理框架，如图 4-2 所示：

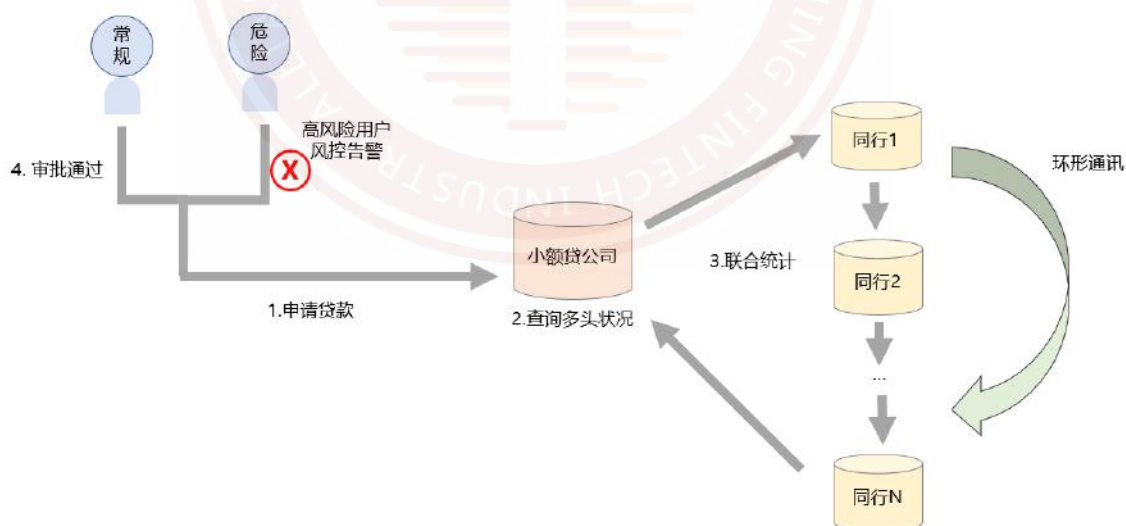


图 4-2 联合统计原理框架

联合统计采用安全求和的方法，隐藏各方真实数据保护各方

数据安全，安全求和的全局过程为环形通信。

基于腾讯云的隐私安全计算平台，小额贷公司可以排查多头借贷和超额借贷的情况，对于高风险用户，可以得到来自平台的高风险报警。因此，小额贷公司可以在防范风险的同时避免各方信息泄漏。

（3）金融机构同运营商数据合作风控

案例一：征信白户增信

主要应用技术：联邦学习

以某银行个人消费贷款申请评分模型为例，该产品的特点是全线上、无抵押，用于满足客户装修、购车、旅游、留学等多方面的用款需求。在风控审批中，该银行可用的数据有：客户在行内留存的个人信​​息以及客户的信用分数据，但如果客户为银行新户时，则没有足够的行内数据可以参考，或客户属于征信白户，即从未办理过贷款业务，也从未申请过信用卡，对于此类客户，很难对其信用水平进行准确评估。针对这类情况，可以引入外部公司数据进行联合建模，本项目利用运营商通话标签数据为客户增信，提升模型的预测能力。如下图 4-3 所示，在进行联合建模前，首先需要找到银行与外部公司的交集客户，例如双方共同的手机号码，通过 PSI 技术保证双方均无法知道合作方的差集客户。

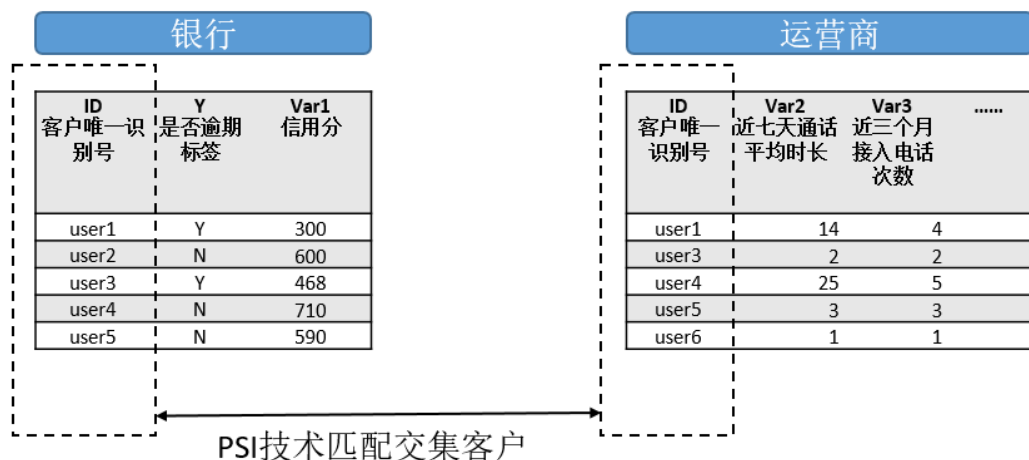


图4-3 PSI技术匹配交集客户

建模时，银行拥有标签数据和征信分数据，运营商拥有通话标签数据，模型训练完成时，双方仅可获得各自对应变量的系数，模型效果相较仅使用自有信用分数据有了显著提高，结果表明，基于联合建模的 AUC 提高了约 10%。利用联邦建模，不仅解决了征信数据来源单一的问题，提高了模型效果，同时也更好地保护了客户隐私，进而帮助银行满足合规要求的前提下，实现了智能风控升级。

（4）金融机构同互联网公司数据合作风控

案例一：信用卡业务反欺诈

主要应用技术：联邦学习

银行信用卡反欺诈一直是业界热点场景，对于黑产数据的沉淀是银行的弱势，而大型互联网公司在欺诈的黑产数据上有较长期的沉淀。工商银行软件开发中心大数据与人工智能实验室与腾

讯集团运用联邦学习技术，在双方数据互不出库、满足用户信息隐私保护以及数据安全的基础上，利用隐私求交技术得有效样本量 4 万多条，通过联邦特征工具加工工行特征数 6 条，腾讯方有效欺诈客户特征特征数 126 条。随后，利用双方特征和联邦树算法进行联合建模训练，得到联邦模型。同时，通过构建本地的三个模型—工行数据模型、腾讯数据模型、工行与腾讯数据联合模型，来验对方数据的可靠性以及联邦学习的稳定性。借助联邦学习技术合规地引入外部数据源，可以获得更真实完整的客户画像，从而大大提高反欺诈预测模型的上限，并完善银行的反欺诈体系，进一步提高银行的风险管理能力。

案例二：长尾客户小额信贷风控

主要应用技术：联邦学习

开放银行作为银行业发展的未来趋势，对银行的数据管理能力、业务创新能力、系统科技水平、风险管理效果都提出了极大的挑战。

先进的信用风险预测模型需要大量数据来提升算法性能，金融行业需要更全面、真实的数据，不仅数据隐私性更高，且各机构间互不信任导致单家机构无法利用大量数据进行模型训练。长尾用户由于缺少信用记录或信用记录较少（例如，申请信用卡、消费贷款等），无法进行有效的风险判断。这一类客户能够获取

的金融服务是有限的，但他们往往是亟需金融服务的人群。

在线上业务自主风控管理体系构建过程中，银行主要面临两方面的挑战：一方面是数据采集范围局限：传统的信贷数据主要基于人行征信数据，在与各合作方开展线上业务合作过程中，可用于线上审批模型及审批策略的基础数据来源较为单一，对于互联网生态环境下的多态数据获取及应用存在较大局限性。另一方面是隐私保护的要求：如何在确保数据隐私的前提下，达成与合作方的数据融合、数据挖掘。

为了解决以上挑战，浦发银行和蚂蚁针对联合贷款业务开发风控模型。双方充分挖掘各自不同的数据价值，提升自主风控能力。多方安全计算技术运用了秘密分享、同态加密等密码学技术，使得数据可用不可见，解决和数据合作中的安全隐私合规痛点。建立更为精准的风控体系，实现差异化风控策略，为长尾客户等提供优质的信贷服务。同时基于模型对业务开展动态管理，有利于快速调整信贷额度，满足用户需求，加速普惠金融的实现。

特别地，对于小微企业开展信贷的时候，由于数据不全，企业经营风险大等原因，小微企业在信贷中往往处于很弱势的位置，在整个信贷领域中，小微企业的信贷占比很小，且还具有融资成本高的问题。但是从国民经济健康发展角度，小微企业是促进就业和市场活力的重要引擎，小微企业对资金流动性具有更高要求，因此该技术对小微企业精准画像，促进小微企业更简单融资，有着巨大意义。

由于安全性上的优势，多方安全计算（MPC）技术是金融业务联合风控场景最为常用的技术，金融机构在进行风控模型建模时，需要通过引入外部数据来增强风控模型效果，以便对小微企业进行更准确的授信判断。

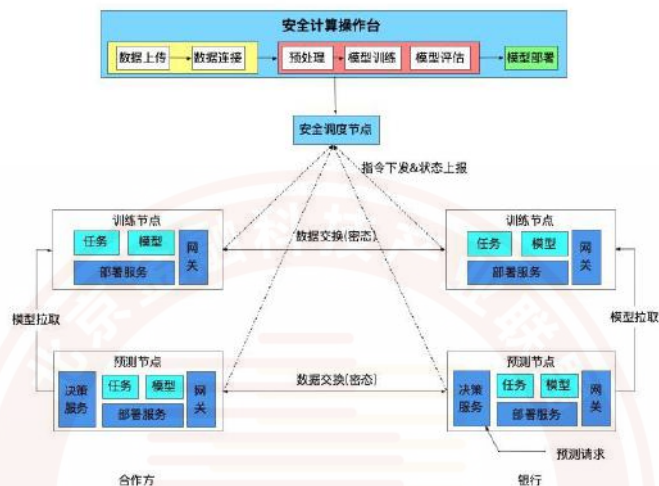


图4-4 训练合作部署架构

图 4-4 所示，为多方安全计算训练合作部署架构。多方安全计算训练流程如下：

步骤一：参与方（银行和数据提供方）将训练数据样本上传至各自的数据存储服务；

步骤二：建模人员在多方安全计算平台通过可视化页面构建数据融合、特征工程、模型训练、模型评估等工作流；

步骤三：工作流以任务形式下发到协调器，后者将任务调度到对应节点的训练引擎；

步骤四：训练引擎根据任务描述，读取本地样本数据；

步骤五：训练引擎使用多方安全计算协议与对端协同完成一次训练任务；

步骤六：训练任务完成后，训练引擎将模型文件保存至各自的模型存储服务。

模型应用主要是预测过程，多方安全计算预测流程（一次预测过程），预测合作部署架构如图 4-5 所示。

步骤一：决策服务作为预测请求统一入口，金融机构从本地发起一次预测请求；

步骤二：决策服务根据模型 ID，请求预测引擎；

步骤三：预测引擎从银行本地的特征服务获取在线特征，利用本侧的模型参数计算预测分数；

步骤四：银行预测引擎向数据提供方预测引擎发起预测请求；

步骤五：数据提供方预测引擎从特征服务获取在线特征，利用本侧的模型参数计算预测分数，并返回给银行；

步骤六：银行预测引擎将两侧预测分数相加得到最终的预测结果。

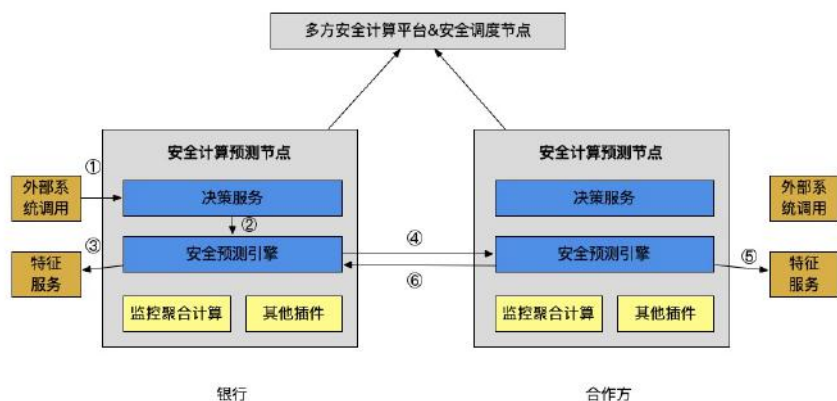


图4-5 预测合作部署架构

通过联合建模有效提升了模型效果，融合双方数据后的模型KS 平均提升 5%以上，业务合作机构实现稳通过率的前提下风险平均下降 20%。该方式在风控场景具有通用意义，不仅提升了风控模型效果，同时还有效保护了数据安全，为跨机构之间持续合作起了标杆作用，该方式也获得更多合作机构的认可。

（5）金融机构同金融科技公司合作进行风控

案例一：黑名单共享

主要应用技术：多方安全计算-隐私求交、隐私信息检索

招商银行在预测行内交易、企业风险等级时，主要使用行内信息，行外尤其是同业银行交易信息及黑名单、风险等级标志信息对招商银行构建反欺诈模型有很高的价值，但因数据隐私安全法律法规的相关规定，同业银行的此类信息无法获取，这为招行训练反欺诈交易模型设置了上限。

基于此，招商银行与平安集团合作，技术侧与平安科技合作，利用隐私求交技术使招行与平安银行安全共享名单数据，实现“数据可用不可见”，优化现有模型，模型结果用于实际业务流程。

如图 4-6 所示，参与双方分别部署隐私求交的查询模块和数据加密模块。参与双方利用匿踪查询的相关技术，通过混淆、去混淆、不经意传输（OT）、验证，在不共享底层数据前提下共享双方名单数据。

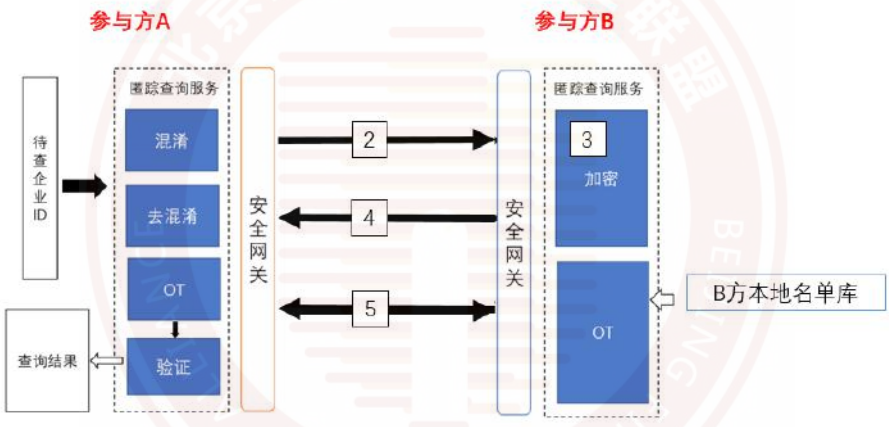


图 4-6 黑名单共享过程示意

该方案可保护双方交集外客户信息安全，使查询发起方无法获取被查询方数据中除查询 ID 以外的信息。被查询方无法获取查询方具体查询的 ID 信息，不知道查询方有没有匹配到的结果，保护查询意图。

依托隐私计算求交技术，可安全有效地使招商银行与同业银行共享名单，获得更真实完整的企业交易信息，提高模型上限。

同时，该案例为同业银行形成了普适的技术解决方案，未来如有更多同业银行加入合作中来，可在反欺诈反洗钱、多头抵押多头贷款等更广阔领域实现业务价值。

（6）金融机构集团内数据合作进行风控

案例一：境内外信用总账管理

主要应用技术：多方安全计算-联合统计

信贷业务作为银行主营业务之一，受到了政府监管机构和社会的重点关注。图 4-7 为部署示意图，农行首先在信贷风控领域应用多方安全计算技术对境内外业务进行有效风险管控。当前，农行信贷管理系统境内和境外应用是隔离的，境内外的客户和用信信息均为分开存储。在境内外联动业务处理中，采用将境外信息汇集到境内进行加工处理的模式。境外业务涉及境内外联动且办理金额较大，亟需实现包括境外数据的客户信用总帐功能，以便在贷前、贷中、贷后业务办理过程中提供决策支持。同时境外区域愈发注重数据保护，在监管趋严环境下，需要探索实践采用技术手段实现数据的安全合规共享。

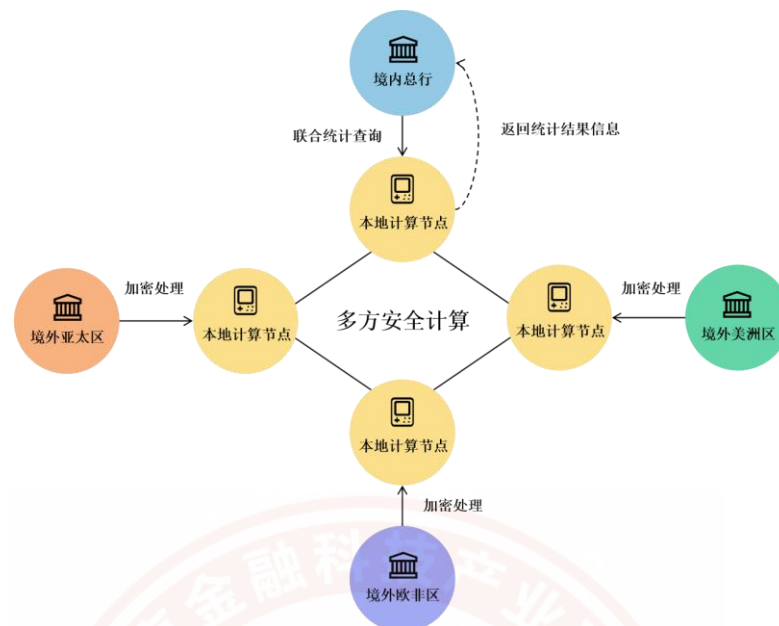
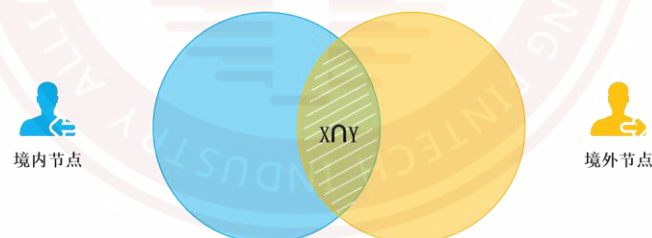


图 4-7 部署示意

如图 4-8 所示，首先利用了隐私集合求交实现对在境内外开展共同业务、存在信用余额的客户进行筛选。

隐私集合求交（PSI）



输入：境内开展业务客户
 $X = \{A, B, C, D, E\}$

输入：境外开展业务客户
 $Y = \{C, D, E, F, G\}$

输出：共有业务客户
 $X \cap Y = \{C, D, E\}$

输出：共有业务客户
 $X \cap Y = \{C, D, E\}$

- 参与方无法获得除结果之外、它方输入中的其他隐私信息
- 输入可根据实际情况，选择参与方同时获得或某一方获得

图4-8 求交找到共同客户

再利用秘密分享的方式实现境内外信用总账计算，如图 4-9 所示。

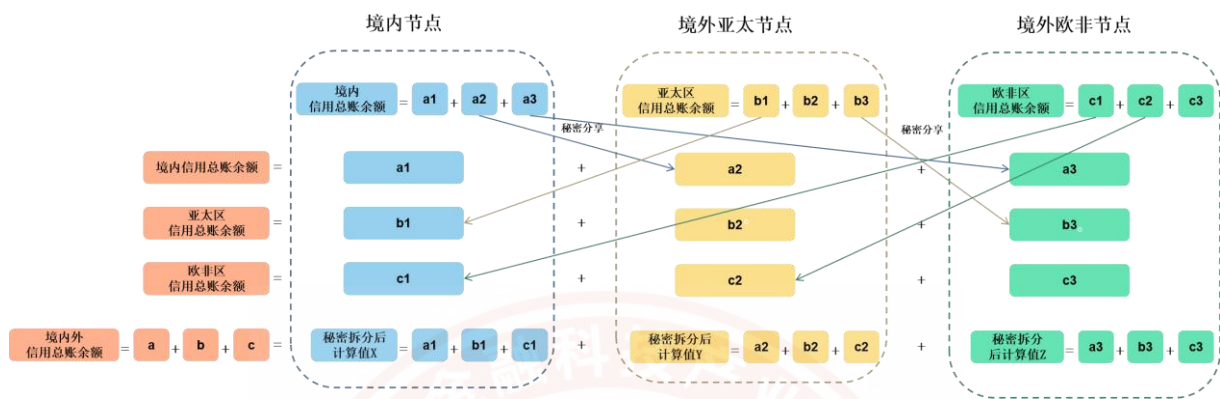


图 4-9 秘密分享计算

用户在信贷管理系统中对其所授权的客户进行境内外信用信息查询时，正式触发多方安全计算，境内外各参与节点进行数据加载、联合计算，并将最终计算结果返回。

利用多方安全计算技术实现客户境内外信用总账管理，避免了敏感数据全量交互，既满足了监管合规要求，又提升了风险管控能力，同时从技术上强化了防护体系，形成了数据保护与共享的刚性约束。本案例实现了客户境内外业务信用情况的全景视图管理，为业务决策和风险管控提供支持，为后续应用场景拓展提供了有益借鉴。

(7) 其他风控解决方案

案例一：违约风险评估

主要应用技术：联邦学习、隐私信息检索

当银行客户向银行提出借贷申请时，银行需要在贷前整合各种方面的资讯来评估违约风险。传统模式下，银行如果需要引入外部数据来提升效果，需要分享明文的还款表现数据给合作数据方用以建立评分卡。这样的做法一定程度上泄漏了客户隐私，也违背了监管要求，因此未能普遍采用。

为了发挥外部数据价值，提升模型预测的准确性，银行现在可以联合其他数据方进行联邦建模，在保护客户隐私数据不泄漏的情况下，利用外部数据建立模型。

基于上述需求，腾讯云隐私计算平台为某银行提供了基于联邦学习技术的金融风控解决方案，其技术框架如图 4-10 所示。

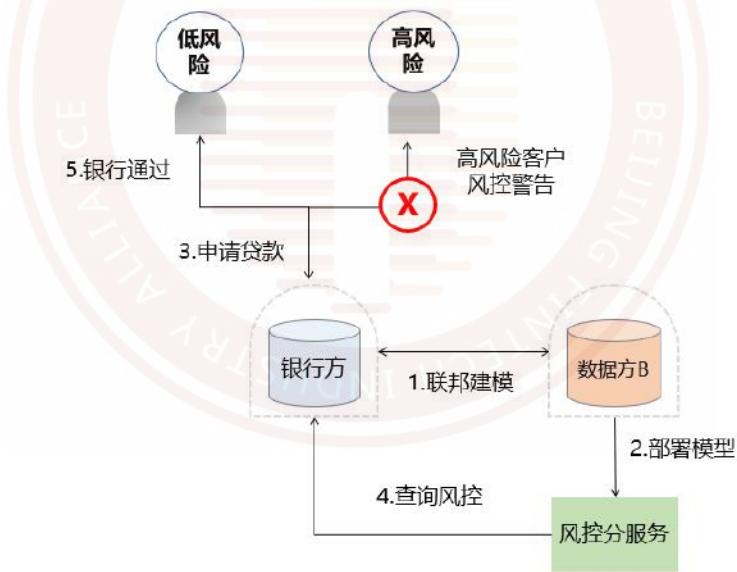


图 4-10 腾讯云隐私安全计算金融风控技术框架

在车贷场景中，当客户向该银行提出车贷申请时，银行将在贷前使用申请评分卡（A 卡）评估用户逾期还款风险，但仅使用人行征信数据和行内积累的数据，对好与坏客户的区分能力有

限。为了提升评分卡的区分能力，银行需要联合其他数据方建立新的评分卡。同时，为了保障客户信息安全，使用了联邦学习技术进行建模。并在实际生产应用时，银行通过匿踪查询服务提供用户借贷逾期风险评分，以避免客户 ID 的泄漏。

在联邦建模中，所有环节都保证银行及数据方的信息安全，包括建模前的特征工程处理，建模过程中采用同态加密技术和代数混淆方法保护中间参数的传输，模型训练后通过匿踪查询的办法来防止客户信息泄露。

同样的模式还可以应用于保险公司的赔付风控等场景，用于解决承保时，保险平台仅调用内部数据和车主历史出险记录，对车主的出险率评估能力有限而导致的出险率过高的问题。

在金融机构的上述金融风控场景中，使用联邦学习技术，可以实现从引流阶段的营销服务到核保、赔付的全面风控服务，保障小额贷、车贷、保险等业务顺利开展。

案例二：小微金融风控

主要应用技术：多种技术融合

小微企业融资难、融资贵是世界性难题。我国一直高度重视小微企业发展融资支持工作，将小微企业作为普惠金融重点服务对象。特别是在新冠肺炎疫情冲击下，有关金融管理部门针对小微企业发布了一系列金融支持措施，要求银行等金融机构发放更多小微企业信用贷款，保障国家就业岗位基本稳定。但由于小微企业融资过程中普遍存在信息不对称问题，其经营风险较大、财

务制度不健全^[16]，金融机构核实小微企业信息成本较高，难以有效评估小微企业贷款风险。

某股份制商业银行与洞见科技进行合作，将行内数据与地方政务数据、外部社会数据（含运营商、航旅、招聘等）融合，运用隐私计算、区块链、大数据等技术构建的智能风控产品，提供安全的数据查询服务、风控数据分析、联合建模、多方数据规则和模型的部署与管理功能，推动解决中小微和个人经营贷金融产品信贷用户的风险评估和决策问题，技术架构如图 4-11 所示。



图 4-11 大数据隐私计算智能风控平台技术架构

在传统的解决方案中，小微金融风控非常依赖大量数据的积累，样本量不足、样本维度不足是最常见的问题和难点。该案例利用同态加密、不经意传输、秘密共享、混淆电路等安全多方计算技术帮助银行构建了大数据隐私计算智能风控平台，将银行的行内信贷用户申请信息、存款、理财、行为偏好等数据和工商、运营商等其他合作方数据进行了安全融合，极大丰富了普惠金融

信贷用户风控数据特征维度，扩大了其他合作方数据开放程度。

通过本案例，帮助该银行联合了包括工商、税务、水电、司法、运营商、征信机构等十余家跨行业数据源提供的上千个数据维度的外部大数据进行中小微和个人经营贷产品的风控，经测算，基于安全多方计算进行训练得到的信用评分模型无损于传统方式得到的模型，甚至 KS 值还比原来高出 2.5%，并且有效优化了银行建模分析决策路径和信贷风控流程，平均审批效率较之前提升了 30%，不仅大幅提升了行方的风险管理水平，而且也极大优化了客户申请体验。

该案例的创新是一种金融机构数字化与智能化转型的安全保护新范式，能够为银行等金融机构基于多方大数据的普惠金融业务提供更安全、更高效、更精准的安全数据智能与隐私保护计算技术赋能。通过将隐私计算、人工智能、大数据等前沿技术有效结合，搭建桥梁打通“数据孤岛”，赋能数据价值，不仅助力银行科技基础平台能力大幅提升，也极大促进了科技创新技术未来在精准营销、反欺诈、反洗钱等其他金融业务应用场景的工程落地和实践。

该案例荣获了中国科技“2021 年度产业数字化优秀案例”、中国电力大数据联盟“2020 年大数据应用十大优秀案例”、江苏省互联网协会“2020 江苏省区块链典型应用案例”。

（二）智能营销场景

近几年结合大数据、人工智能等技术描绘用户画像的精准营销方式已在金融行业内广泛开展，整合多机构间多维度的数据，构建更立体的用户画像，是达到资源优势互补、开拓市场广度和挖掘服务深度等营销目标的有效手段之一。但金融机构间的用户资产画像信息相互割裂，且金融数据涉及用户隐私，具有一定敏感性，也是金融机构的核心资产之一，具有较高商业价值，难以以明文形式共享使用。隐私计算技术可以帮助金融机构以密态方式共享各自的用户数据进行营销模型计算，根据建模结果制订营销策略，实现联合营销目的。

（1）金融机构同运营商数据合作营销

案例一：理财产品客户营销

主要应用技术：隐私求交

某区域性银行希望在存量的信贷客户中找到有理财产品需求的客户，开展交叉销售。但对于存量信贷客户缺乏全面的洞察，特别是客户理财兴趣类的标签比较稀疏。该银行希望在自有数据不出库的情况下，引入第三方数据，更好地刻画用户的金融属性和理财产品偏好，提升营销效果。

电信运营商拥有丰富的手机使用行为的标签，比如：手机上安装理财 APP 的个数、每月登录理财类 APP 的次数、每月理财类 APP 累计使用时长。这些信息对于判断用户对理财产品的兴趣度

是很有价值的。



图4-12 合作双方求交客户

用传统的 MD5 加密的方式输出电话号码给第三方数据源时可能带来大量的用户号码泄露，不符合数据安全合规要求。该银行与当地运营商开展合作，如图 4-12 所示，在保证银行和运营商的原始数据不出域的情况下，基于 PSI 安全求交集找到共有用户群，通过标签的组合筛选找到对理财产品有兴趣的目标用户，然后由银行自有坐席进行外呼营销。

如图 4-13 所示，该模式下运营商和银行的原始数据不出库、非共有 ID 不被透露，比传统的 MD5 方式更安全合规，同时也保持了运营商对数据的控制权。银行基于运营商数据丰富了自身存量客户的标签与画像，提升了营销精准度和转化率。

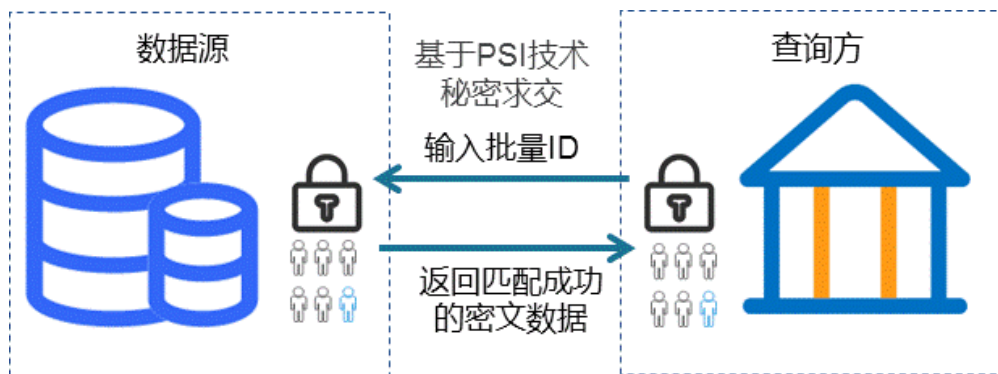


图4-13 合作示意图

案例二：保险行业存量营销

主要应用技术：联邦学习

大部分传统型保险公司，面临着激烈的市场竞争环境，因运营成本高，处于利润稀薄，甚至承保亏损的困境。随着大数据、人工智能技术的发展，保险公司开始更多运用数字化手段解决自身运营问题。保险公司和运营商基于以下原因有普遍的数据协作需求：保险是高频的复购场景，保险公司有大量的存量沉默用户和新增赠险获客，存在多种存量用户挖掘的需求。在保险公司和运营商的合作中，保险公司通过关联运营商数据的身份证号—手机号关联图和用户画像标签，建立了保险意向人群的定位和挖掘模型。

原合作存在多个数据隐私保护层面的痛点。首先，在数据交互形式上，由于原有技术手段仅能通过 MD5 哈希加密算法隐藏用户的真实 ID（本合作中为身份证号），而增加了数据被碰撞攻击，业务留存和反向查询的风险。其次，由于《网络安全法》和保险公司自身数据安全政策的限制，用户挖掘依赖基于业务经验

的实时策略调整，保险公司无法向运营商提供历史的成功购险转化名单用于构建高质量的机器学习模型，保险意向人群的挖掘精度尚无法充分发掘数据的价值。数牍科技为保险公司和运营商提供了基于隐私计算的数据协作方案，有效解决了以上双方数据交互的痛点。

一是通过隐私保护集合求交技术对双方用户 ID（本合作中为身份证号）的隐私保护。匹配后仅保险公司可知道匹配的用户，运营商无法留存或反推出用户身份，仅当保险公司明确了需对某用户建立触达后，双方完成 ID 信息的交互，解决了保险公司的 ID 泄露顾虑。二是通过联邦学习技术，成功将保险公司的历史成功投保标签和运营商数据的用户画像标签结合到一起，建立了基于纵向联邦学习的保险意向模型，在双方不交互原始数据的前提下，实现了数据价值的交叉挖掘。

某保险公司通过使用运营商的用户画像标签建立挖掘策略，恢复联系的用户综合投保率约 4.2%。通过隐私计算技术建立的隐私购险兴趣挖掘模型，在对既有人群数据中的评估中，模型评分前 50% 人群购险率达到了原有表现的 1.8 倍，见图 4-14。

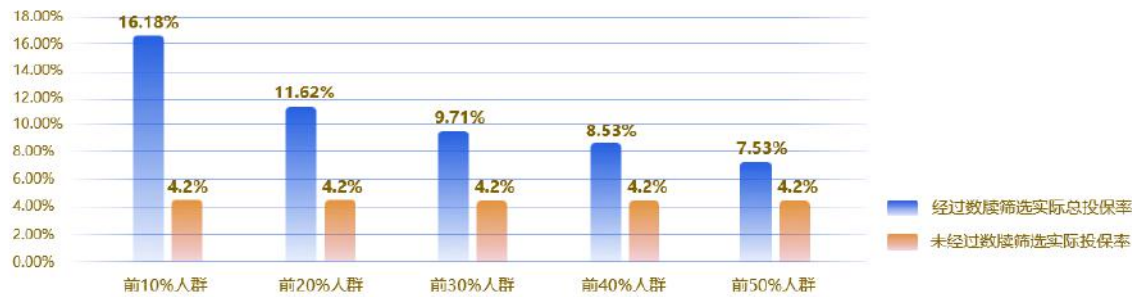


图 4-14 运营商—数牍—保险公司隐私数据用户意向识别结果图

在此案例中，科技公司结合自主研发的隐私计算技术体系，运用联邦学习等隐私计算技术，帮助保险行业客户实现保险合作企业与保险公司的底层数据支撑业务打通服务，在保护用户隐私及数据方用户信息可用不可见的前提下，拓宽数据样本量及数据维度，帮助保险客户实现存量客户高效触达，支持客户输出画像标签的精准预测，极大地提高了用户转化率。

（2）金融机构同政务数据合作营销

案例一：批量筛客和流量拓客

主要应用技术：联邦学习

政务场景是富含“数据金矿”的核心场景。为挖掘数据“金矿”，获得企业客户价值实现精准营销，利用纵向联邦学习技术，招商银行与政务部门合作，分别部署联邦节点，利用行内和政府数据训练联邦模型之后，无需各数据方披露底层数据即可使用模型预测，模型结果用于各种实际业务流程。同时部分项目根据实际情况利用区块链技术将各数据方行为数据上链，方便追溯验证监督以及出现问题之后追责。

招商银行在营销拓客场景已在深圳、北京、厦门、杭州等多地分行实现中小微企业场景落地，全国范围合作政务部门十余家。

在联邦获客营销问题上主要采用批量筛客、流量拓客两类解决方案。联邦获客营销解决方案如图 4-15 所示。

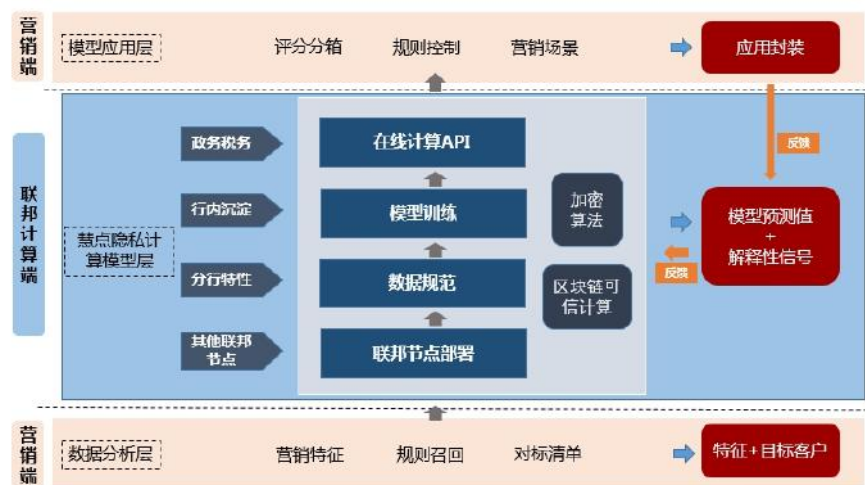


图4-15 联邦获客营销解决方案

（a）批量筛客

模型训练完成之后，利用行内客户名单批量预测，识别不同领域优质客户，不同客户有针对营销贷款、存款等不同业务。具体包括 3 个阶段：

模型训练阶段（无需授权）：利用隐私求交算法（PSI），找到双方用户交集，明确双方训练数据特征维度。

模型预测阶段：该过程一般需要用户授权。一般的做法是先由招行根据自有数据筛选出行内开户客户，但还未办理各种业务的潜在营销客户名单，利用模型预测，将客户分至各个业务领域的高价值客户。

批量拓客阶段：名单筛选完毕进入线上线下营销阶段，将不同客户推荐至其对应高价值领域，“千人千面”，形成定向营销。

（c）流量拓客

模型训练过程与批量筛客相同，但模型预测过程不同。招商

银行信贷产品利用数据合作方政务门户网站用于引流，单个企业点击申请授权，通过联邦实现线上拓客，让企业不再等“贷”，有效助力普惠金融场景，解决中小微企业融资难、融资贵问题。

依托隐私计算技术，招商银行与多政务部门继续深化合作，共同挖掘数据价值。一是合作研发企业信用评分模型，结合银行客户营销服务需求进行个性化定制，进一步丰富完善我行企业用户画像，为精准营销提供有力支撑。二是合作信用数据校验，对于开户、授信环节需要人工审核校验的一些隐私数据，通过联邦学习技术进行校验，进而实现流程优化，提升客户体验。三是利用政务门户网站流量，为招行进行获客引流，同时也为政务部门监管提供助力。

（3）金融机构集团内数据合作营销

案例一：工商银行集团内联邦助力保险营销

主要应用技术：联邦学习

工商银行软件开发中心大数据与人工智能实验室与工银安盛运用联邦学习技术，在双方数据互不出库、满足用户信息隐私保护以及数据安全的基础上，在保险营销场景中完成联合建模。工银安盛销售渠道主要为自销渠道和银保通渠道，和我行用户数据相对隔离。数据显示银保通渠道成功销售客群平均资产比保险自销渠道高很多，该方案拟通过联邦学习下探银保通渠道低资产客群。

在实施过程中，银行共提供了 175 个建模变量，包含客户基本信息变量，如年龄、性别等，和投资类、保险类相关变量。保险特征可用特征约 281 维，包含保险各类客户标识，历史购买保险信息，和部分保险画像特征，如保额、保费、缴费期限、被保险人等。根据场景是否为首购以及客户是否为代发客群两个条件，本方案建立了由 3 个联邦模型组成的集成模型。首购场景仅使用了银行特征，复购转化场景使用了银行和安盛两方的特征。结合联邦建模进行客户探查，可以全面挖掘客户的投保能力，深入细化客户的投保偏好，从而提高保险的精准营销能力，为集团内营销体系提供有力支撑。

案例二：建设银行集团一体化，速盈客群价值提升

主要应用技术：联邦学习

2020 年下半年，由中国建设银行上海大数据智慧中心牵头，建信金科提供技术支持，联合建行集团子公司建信基金，利用隐私计算技术进行了集团一体化建模探索，参与联合建模的双方在生产环境中验证了隐私计算技术在金融产品智能营销场景的可行性。

在本项目中，双方在生产环境中利用真实业务数据进行了基于用户特征维度拼接的纵向联合建模，针对“速盈客群价值提升场景”，实现跨双方隐私计算纵向模型建立，定位目标客群，助力建信基金侧的客户价值挖掘和提升。

在项目的实施过程中，需要利用存储在总行生产环境中的客

户数据与云环境外的建信基金产品数据进行纵向联合建模，由于建行云环境的网络连接安全性要求，双方最终通过在 DMZ 区部署转发路由服务实现，部署示意如图 4-16 所示。



图 4-16 部署示意

在具体的模型建立过程中，双方通过纵向联邦学习的方式，综合客户基本信息、客户金融属性、产品特征等几十维数据特征，针对用户对产品的偏好情况进行模型训练及预测。在实际的营销中，针对评分前 5%客群的响应率相对于单边模型提升 34%。

该项目实现了集团内不同机构间数据融合建模功能，在确保双方原始数据不出域的前提下获得模型结果，实现业务价值提升。

该项目被中国信息通信研究院和中国信息标准化协会大数据技术标准推进委员会认定为“隐私计算优秀案例”，获得“星河”奖。

（4）金融机构同互联网公司数据合作营销

案例一：线上定向广告隐私计算解决方案

主要应用技术：联邦学习

在线广告产业上下游包括广告需求方、媒介方以及广告信息交换平台^[17]。在很长一段时间内，大量互联网头部公司在实践中获得最大规模化营收的主要来源即是在线广告。国外诸如 Google, Facebook, 国内如腾讯, 阿里巴巴, 今日头条, 百度等公司, 广告营收均占据较高的份额。在需求端, 各个公司或品牌方每年投入大量的预算在线上广告中, 来达到推广公司产品, 获取未来营收等目的, 如银行希望在网络上找到本行信用卡产品的潜在用户群, 并通过广告争取这批潜在用户。当宏观经济增速放缓, 商家的营收能力开始下滑时, 将对应地削减广告预算, 同时要求更高的 ROI (Return on Investment), 对具有更高效率、更低成本的智慧广告提出了迫切的需求。

线上广告核心技术包括用户定向及竞价等。基于联邦学习的广告技术, 将有助于降低用户拉新的成本。利用联邦迁移学习, 能够更好地整合多方数据, 构建用户洞察与定向策略。在广告中重要的技术环节 RTA (Real-time API) 中, 有助于广告主高效获客, 同时避免重复投放的资源浪费。通过在 RTA 中引入联邦技术, 有助于进一步提高双方数据的安全性, 同时有效降低成本。在安全性方面, 通过联邦学习技术, 广告主与媒体双方的用户特征与标签随机加密, 双方无法知晓用户个体是否发生后端转化以及具体风险标签。同时联邦学习中差分隐私技术的接入, 将混淆与打算双方数据, 使对方或第三方仅能看到整体情况, 而无法识

别数据中的任何个人。微众银行的联邦学习联邦广告解决方案，在数据安全和隐私的保障下，更有效的连接需求方和供给方，提升转化率，降低成本。

（5）其他营销解决方案

案例一：保险营销业务的评估与决策

主要应用技术：多种技术融合

随着大数据时代的快速发展，数据的价值得以彰显，保险公司在发展过程中积累了大量客户相关数据，但这些内部数据或是纸质材料没有形成结构化数据，或是没有被充分利用起来进行挖掘分析，而且数据多集中在保险这个垂直领域，难以对客户进行精准的保险产品营销和定价，因而保险公司往往需要结合自有的存量客户数据和其他领域的外部数据联合建模，综合对其客户的消费行为、需求偏好和出险风险等进行评估分析，从而得到更加精准的有保险需求的客群和出险概率、赔付成本等预测结果。

某保险公司与洞见科技合作，基于隐私计算、区块链、智能应用等方面的技术，建立了贯穿保险产品营销筛选、产品定价、效果分析、模型优化等全生命周期的大数据隐私计算智能保险营销平台，从安全融合到智能发现，再到决策应用和数据上链，全面提升了多方大数据在保险行业的市场营销应用价值和效率，业务应用流程如图 4-17 所示。

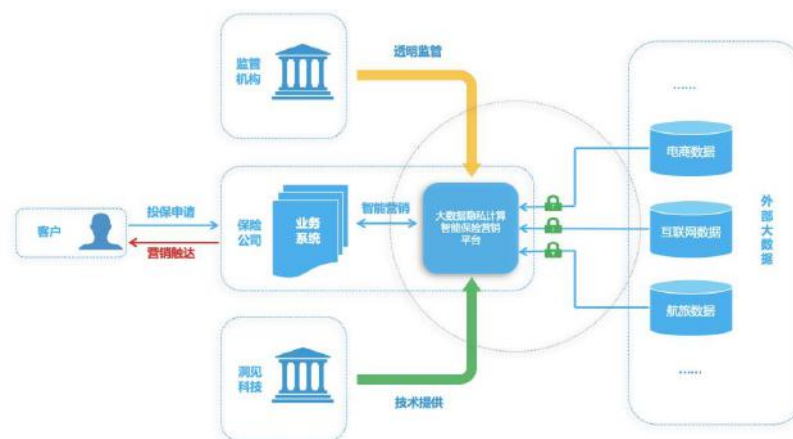


图 4-17 大数据隐私计算智能保险营销平台业务应用流程

通过本案例有效帮助了保险公司判断风险、预测出险概率和赔付成本，达到了降本增效的目的，在试运行期间，该保险公司的医疗险、健康险和车险等业务均显著提高了转化效率，节省了大量人工客服触达的时间和成本，不仅大幅提升了保险公司的精准营销管理水平，而且极大优化了客户投保体验。

本案例通过安全多方计算技术引入更多外部大数据作为保险垂直场景数据的有益补充，完善用户画像，精准挖掘潜在客户，实现了个性化保险定价，既提升了保险公司市场营销的精准程度，更保护了营销模型和规则的商业机密，这种行业示范作用无疑给其他保险公司的市场营销业务也带来了新的引领。

在监管合规方面，保险业向来面临强监管，该案例基于隐私计算的大数据精准营销帮助保险公司和合作机构方在不暴露各自客户隐私数据的前提下，安全可信地实现了保险营销业务的评估与决策，满足国家监管政策、隐私信息保护法律法规的要求，以安全合规的方式进行数据流通与协作。

在社会价值方面，试运行期间保费贡献规模超过了数千万元，切实为广大人民群众提供优质的医疗险、健康险和车险等服务做出了贡献。案例的成功落地也预示着隐私计算未来在保险等金融机构之外，还可以为医疗、公安、集团、政府等社会各业提供隐私计算、数据治理、数据安全等各类解决方案，助力国家实现安全可信金融和政务数据智能升级。

该案例荣获中国信通院 2020 大数据“星河”隐私计算优秀案例。

（三）智能运营场景

通过联邦学习、多方安全计算、隐私信息检索等隐私计算手段安全利用第三方数据，可实现客户分群、客户特征描述、产品需求偏好分析等，形成对客户的全维度画像，充分发掘客户与金融机构间的关联行为以及需求状况，进而为特定群体客户、特定业务引流，提升促活效率。主要在金融机构集团内数据合作进行运营。

案例一：企业级全流程智能运营

主要应用技术：联邦学习

某银行与瑞莱智慧合作建立智能运营全流程把控，客户遇到的问题是在不同地市有不同的分行难以实现互联互通，如何对客户进行精细化管理细分客户群体，针对不同类型的客户制定不同的业务策略，改善服务，提高效率，降低成本。

集团内部各个子公司部署隐私保护机器学习节点，实现内部以客户为中心的信息互联。同外部高价值客户特征数据源利用隐私保护机器学习平台对接，对客户特征进行建模交叉匹配，精准圈定目标客群。

案例将银行内部存量数据信息打通，补全用户画像，同时运营转化率整体提升 20%。极大的提升了用户在运营场景的效率，为用户创造了价值。

案例二：汽车金融集团内数据监管运营

主要应用技术：联邦学习

国内某汽车集团正在做集团内的数字化转型项目，其中一个重要的阶段是“融资租赁去担保化”，其旗下的汽车金融事业部承担了该阶段的重要角色。但是在实施过程中发现，集团内部因不同子公司主体不同，无法使用原始数据进行共享，进而导致客户的维度不齐，去担保化存在障碍。

解决方案如下。首先，汽车金融事业部明确了集团内的数据来源，确定与集团内的重机事业部进行联合建模；其次，汽车金融事业部协同同盾科技，根据同盾科技的业务经验，对双方需要提供的数据类型进行大概的梳理；最终，根据场景明确汽车金融事业部数据（如标签信息、账龄信息、滚动率信息、客户基础信息等），同时也对重机事业部数据提出了需求（如设备使用信息、位置信息、维修信息等）。

再结合集团内的数字化转型的要求，明确汽车金融事业部的

“去担保化”目标——汽车金融的贷中后场景。本项目采用联邦学习的方式进行模型的构建，技术和解决方案由同盾科技提供，具体实施步骤如下：

第一步：在参与双方部署同盾的智邦平台，部署的前提是，双方均准备好部署环境；与此同时，同盾科技根据自己多年的金融业务经验，对双方的数据分别进行探查工作，主要包括有：样本定义、数据清洗（因重工行业数据一般都比较混乱）、关键特征衍生等操作。

第二步：待联邦平台部署完成后，将汽车金融事业部的样本及字段信息、重机事业部的字段信息分别上传自本地的联邦学习平台，并选择合适的算法进行训练；为满足模型的可解释性要求，本次选用的是逻辑回归算法，建模样本万余个，建模字段百余个。

第三步：对模型参数进行调优，并输出最终的评分结果和分数分布，并根据模型可解释性要求，对模型进行评审。

第四步：将通过评审的模型进行分布式部署，并通过加密的方式进行结果的调用。

值得一提的是，本项目集团总部提出了监管的需求，即对数据的使用，模型的调用等需要总部进行监管，因此，采用的是同盾科技“弱中心化”的联邦学习方案（并结合区块链技术），保证业务顺利进行的同时，集团对整体数据的使用过程都有监管、监控、可追溯的权利。解决方案示意图如图 4-18 所示。

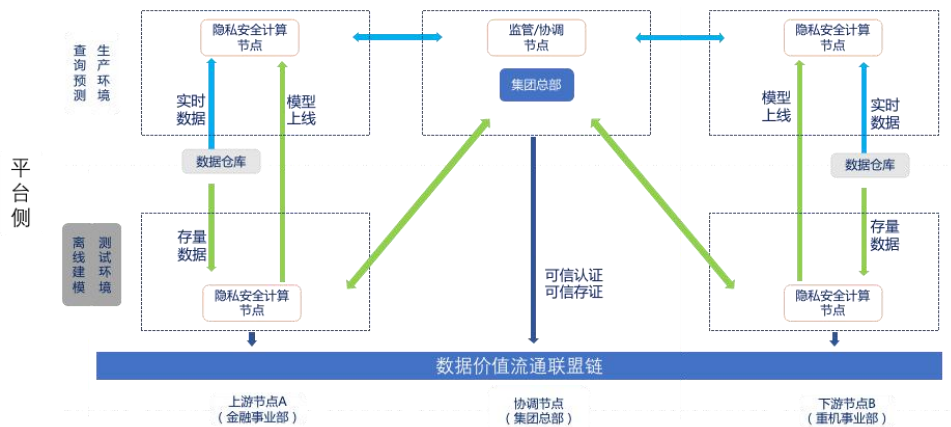


图4-18 解决方案示意

经验证本项目在模型效果和实际使用的过程中，都取得了不错的效果。构建的模型 $KS=0.33$ ， $AUC=0.87$ ，模型 $PSI=0.006$ ，符合汽车金融事业部对上线模型要求；在实际调用过程中，分数分布与建模样本分数分布基本吻合，模型表现效果良好。统计结果如图 4-19 所示：



图 4-19 统计结果

本案例旨在通过联邦学习的方式，将集团内的数据赋能给金融事业部，一方面，应用联邦学习技术充分激活内部数据价值，解决数据孤岛问题；另一方面，去担保化有助于放量，帮助业务拓展。实现集团内数字化转型的关键一步。

（四）隐私信息检索场景

隐私信息检索，又可称为匿踪隐私查询、匿踪查询，在保护查询方不泄露查询 ID 信息的前提下实现数据查询。可以单独使用该技术在某些金融场景，也可以结合多方安全计算、联邦学习等技术一起使用，达到最大限度保护信息的目的。

案例一：客户回捞

主要应用技术：隐私信息检索

某银行与瑞莱智慧合作建立拒客回捞算法，运用匿踪隐私查询技术进行数据的隐藏。前期客户个人线上贷款产品率仅为 10%，需要从拒客中回捞优势客户，同时具备高价值客户信息方存在泄漏被查询者隐私的风险。

方案应用匿踪查询查询外部敏感数据，评估客户的真实信用情况，再运用半监督学习、迁移学习解决幸存者偏差问题，判断之前被拒绝客户是否为优质客户。

通过匿踪查询的方式，外部数据源机构获取不到精确的查询 ID，保障了信息安全。从 250 万的拒绝客户中回捞 10% 优质客户，结果表明：通过率提升 50% 而且逾期率保持不变，极大地提升了

用户客户回捞的效率，为企业和用户创造了价值。

（五）供应链金融场景

传统供应链金融模式下，由于信息不够透明导致银行可融资主体范围过窄，主要依赖核心企业的控货和销售能力^[18]，由于其他环节的信息不够透明，银行出于风控考虑往往仅愿意对上游一级供应商提供应收账款保理业务，或对其下游一级经销商提供预付款或存货融资，而对融资需求迫切的二三级供应商和经销商无法得到金融支持，供应链金融难以大规模开展。

通过隐私计算可以实现上下游间数据的安全联合统计分析，在数据可用不可见前提下，切实保护企业隐私，打通企业税务、电力、交易流水、票务数据等多源数据，构建供应链上下游信息对称共享体系，扩大供应链金融整体服务半径。

利用隐私计算技术通过将其他领域数据与金融数据的融合，为银行小微企业供应链融资中风险防控等场景的模型补充数据内容，增强小微企业信用评估能力，扩大银行普惠金融服务半径，切实提升小微企业获得融资服务的效率。

将电信运营商数据、电力数据、税务数据、政务数据、票务数据、企业内部的 ERP、MES 数据，结合银行征信数据，用于小微企业反欺诈、贷后预警、企业黑名单等模型建设。在各方原始数据不出本地的前提下进行融合应用，打破机构间的数据融合壁垒，提高各方数据融合应用的效率。

案例一：供应链金融中的风险监测

主要应用技术：联邦学习

电信运营商拥有上亿级的真实用户，数据具有规模大、来源广、种类多、实效高等特点。用户特征包含用户基本信息、用户终端信息、通话及流量信息以及用户 APP 使用统计信息等。

传统的联合建模方法不仅无法满足双方数据隐私保护的要求，同时也无法将各方的数据资源最大化利用。一类合作方法是需要银行提供用户 ID 及数据标签等到电信运营商处进行建模，建模完成后，银行通过接口调用模型进行预测。这种方法在数据合规要求下，需要将数据脱敏，影响准确率，同时也存在泄漏用户数据的风险；另一种方法是电信运营商直接提供评估结果，如用户评分等。银行方将运营商的评估结果作为数据标签加入自己的数据集后进行再训练，这种方法未能打破数据孤岛问题，联合建模的效果仍不够理想。

浙商银行与某电信运营商利用联邦学习的隐私计算技术，在保护用户数据隐私、原始数据不出本地的前提下，合法合规地从多源数据中训练出更加准确的联合风控模型，用于小微企业信贷风控建模，为供应链融资中风险防控等场景的模型补充数据内容，打造用于供应链金融的动态风险监控产品，进一步增强小微企业信用评估能力，扩大银行普惠金融服务半径，切实提升小微企业获得融资服务的效率。

方案基于隐私求交将双方数据集根据样本 ID 对齐，然后以

对齐后的训练数据集为基础对逻辑回归模型和 XGBoost 模型进行联合训练，最后根据测试数据集来衡量联合模型的表现。

通过与电信运营商的合作，在各方原始数据不出本地的前提下进行融合应用，打破机构间的数据融合壁垒，提高各方数据融合应用的效率。

（六）反洗钱场景

近年来，不法分子依托人工智能等高科技手段将非法收入清洗为合法财产并融入经济体系的案例屡见不鲜，洗钱行为呈现出更隐蔽、成本更低的趋势，以专家规则、数据排查等传统方式的反洗钱手段难以有效挖掘由新兴技术手段支撑的洗钱行为，造成全球范围内的洗钱数额居高不下。据统计，平均每年全球洗钱总额高达 8000 亿至 2 万亿美元，约占全球 GDP 的 2%到 5%。洗钱活动给社会带来了巨大的负外部性，不仅会造成涉事金融机构信誉受损、对资金流动的监管难度加大等问题，还会助推犯罪活动，影响社会安定，造成营商环境恶化，极大阻碍了社会经济发展。

作为反洗钱金融行动特别工作组 (Financial Action Task Force on Money Laundering, FATF) 的成员国之一，我国金融监管机构积极推进了一系列反洗钱工作举措。特别是在处罚力度方面，我国 2017-2019 年间反洗钱行政处罚总金额逐年上升，且增速加快（2017—2019 年反洗钱处罚金额见表 4-1）。2020 年，我国反洗钱行政处罚总额为 2019 年全年的三倍，涉及银行、保

险、证券、支付、期货、信托、财务公司七类机构。同时，随着《中华人民共和国反洗钱法》和《刑法》等法律规章制度的修订及建设工作陆续启动，我国反洗钱法律体系逐步健全，为加强执法和监管力度打造底线机制。

表 4-1 2017—2019 中国人民银行反洗钱处罚金额^[19]

年份	总处罚金额(亿元)	总处罚金额(亿\$)	总罚金/GDP 占比(亿\$)
2017	1.34	0.198	1.289
2018	1.66	0.251	1.549
2019	2.15	0.312	1.877

然而，面对海量金融交易数据，机构应用的基于规则和名单的反洗钱机制过多依赖历史经验，无法自发捕捉新型洗钱特征和实体关联关系，导致反洗钱成本高昂。在洗钱风险复杂化和国际反洗钱要求趋严的双重压力下，我国金融领域开始积极探索并推动反洗钱机制由传统的基于规则的模式转型为新兴技术赋能的新模式。

依托基于人工智能、大数据等新兴技术的反洗钱监管手段，金融机构能够对海量、多维度的客户数据进行分析，准确识别客户风险及可疑洗钱活动。特别地，在洗钱活动的放置、培植和融合阶段中，资金流动通常涉及多家金融机构（包括银行、信托、证券、保险、支付平台等）及非金融机构（包括拍卖行、工商企业等）。金融机构及监管部门应用人工智能技术对这些数据进行处理分析，能够全方位了解账户关联、交易时序及资金流向情况，

精准侦查可疑行为。然而，受限于数据安全和资源优势的顾虑，金融机构当前普遍拒绝开放共享自身数据，对人工智能技术在反洗钱领域的应用造成数据瓶颈。在此背景下，金融行业亟待依托技术手段打破数据壁垒、实现数据安全共享，为下阶段构建基于人工智能和大数据的反洗钱监管体系夯实数据基础。

由于行业信息不共享，金融机构无法协同分析不同账户的关联关系和资金链路，因而难以有效识别洗钱行为。如此一来，金融机构向监管方上报的黑名单往往误报率过高。据业内人士分析，我国银行业的反洗钱误报率高达 95%^[20]。这便导致监管方在投入大量分析成本后，监管效率依然低下。

利用基于多方计算技术的行业级数据融合平台，金融及非金融机构可构建完善的跨机构反洗钱机制，通过安全共享相关数据进行全链路的资金流动和人物关联分析，实现对可疑洗钱分子的精准识别，降低样本误报比例，减少监管方成本负担。

具体而言，该任务涉及数据输入方、计算方和结果获取方三个角色。其中，计算方为联合建模平台；数据输入方为银行、支付、保险、证券、期货、信托、财务公司等金融机构和工商企业和拍卖行等非金融公司；结果获得方为监管机构。此处例举平台执行基于秘密共享的联合建模功能，安全融合客户及其资金流向的相关数据，联合建立洗钱风险评估模型。具体流程如下：

一是数据输入环节：各数据提供方将客户信息及其资金相关数据通过计算节点以密文形式输入平台。

二是计算环节：平台对输入因子进行预处理后分发至各计算节点，确保每个计算节点仅能获取到单个输入节点输入因子的部分“切片”。各计算节点完成各自计算后将结果提交至计算引擎，由多方计算引擎将计算结果切片后分发给若干个计算节点（每个计算节点仅能收到部分切片，即输出因子）。

三是结果输出环节：平台将结果输出至监管机构并解密为明文结果，即客户洗钱风险评分，用于筛选可疑洗钱分子。随后各计算节点销毁计算因子、计算中间结果和最终结果数据。上述流程如图 4-20 所示。

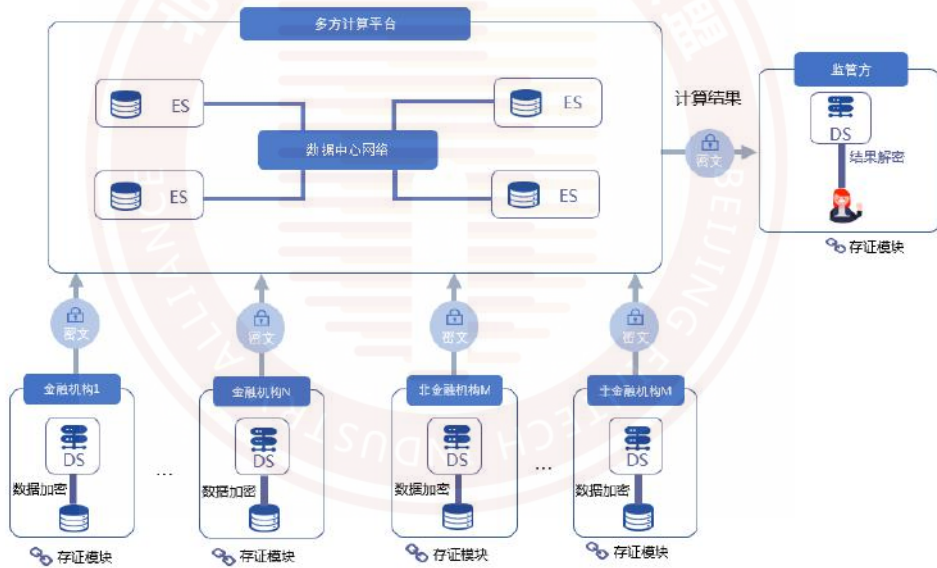


图4-20 流程示意

基于多方计算技术进行数据安全融合应用，金融机构实现了各方的数据“可用不可见，可控可计量”^[21]，同时全流程可验证、可追溯、可解释、可审计、可监管，对金融行业监管合规、监管部门防控洗钱风险均有积极意义，有助于促进我国金融稳定、优

化营商环境、推动经济发展。金融机构方面，应用基于多方计算技术的行业级数据安全融合平台，实现了对客户背景、交易性质、主体关联关系、交易时序及资金流向的全链路监测和分析，精准挖掘具有反洗钱嫌疑的异常资金结构和反洗钱团伙，降低监管合规成本；监管方面，利用该平台，监管部门获得更精准的洗钱风险评级信息，无需分析海量可疑样本数据，从而减轻监管成本，提升反洗钱效率；社会方面，监管机构通过应用智能化反洗钱体系有效打击了洗钱犯罪行为，维护我国金融体系的安全稳定，提升国家形象，促进营商环境优化和资本流入，助力经济发展。

（七）企业级数据流通交易

诸多已有平台对数据交易的理解仅仅停留在原始数据明文共享的阶段。受限于明文数据“责、权、利”难以厘清的问题，这些平台并未实现真正意义上的数据要素流通。例如在政务数据方面，许多地方政府虽然设立了大数据共享中心，但由于各政府部门间数据共享的权责界定不清，担心数据开放会导致自身失去数据优势，导致政务数据的开放共享难以推动。

上述数据交易瓶颈在于明文数据复制成本低、使用没有排他性的特点。具体地，明文数据几乎可以无限地被复制，且可同时被多方使用，其用途和用量难以得到规范和限制。这些特点往往被称为明文数据的优势，但亦是明文数据大规模分享的掣肘。因此在现实中，明文数据一旦泄露便导致数据滥用问题，数据所有

方在分享数据时顾虑重重，尤其是在个人信息及隐私保护要求日趋严格的背景下，不愿、不能、不敢共享数据成为金融、政务、医疗等诸多行业数据开放最大的瓶颈。

在数据要素化战略和数据隐私保护要求并行的时代背景下，隐私计算技术凭借其能够实现数据“可用不可见，可控可计量”的特点，成为推动各行业数据安全合规流通的有效解决方案。

案例一：金控集团跨机构数据统计

主要应用技术：多方安全计算-联合统计

大型金融控股集团中，各金融企业的用户信息常常是分散的。这些用户信息可能存在重叠部分，也就是说不同金融企业之间拥有共同的客户。对不同企业间用户信息进行统计，有助于挖掘更多数据价值。对于跨机构数据统计问题，在传统的方案中，金控集团通常会建立一个大型的数据中心，各金融企业将数据上传至数据中心，最终由数据中心对汇总后的数据进行统计。但是随着社会对用户隐私问题的重视程度逐渐提升，同时由于金融行业的特殊性，各级立法和监管机构出台多项法律法规和监管规定，加强对个人金融数据隐私的保护力度，传统的跨机构统计方法已无法满足对个人金融数据隐私保护的监管要求。如何保证数据传输的安全性和可靠性，如何管理和审计涉及多方交互的数据，并在合法合规的前提下实现跨机构数据统计，成为一个重要技术难题。

在光大集团的联邦数据治理实践过程中，针对跨机构用户资

产求和这一场景，实现基于可验证秘密共享(verifiable secret sharing)的安全多方隐私求和方案，如图 4-21 所示，能够在数据不出本地的情况下，对用户在多个机构的数据求和。该方案在保障光大用户信息绝对安全的前提下，实现了数据的协同计算，最大化地释放了数据要素的价值。

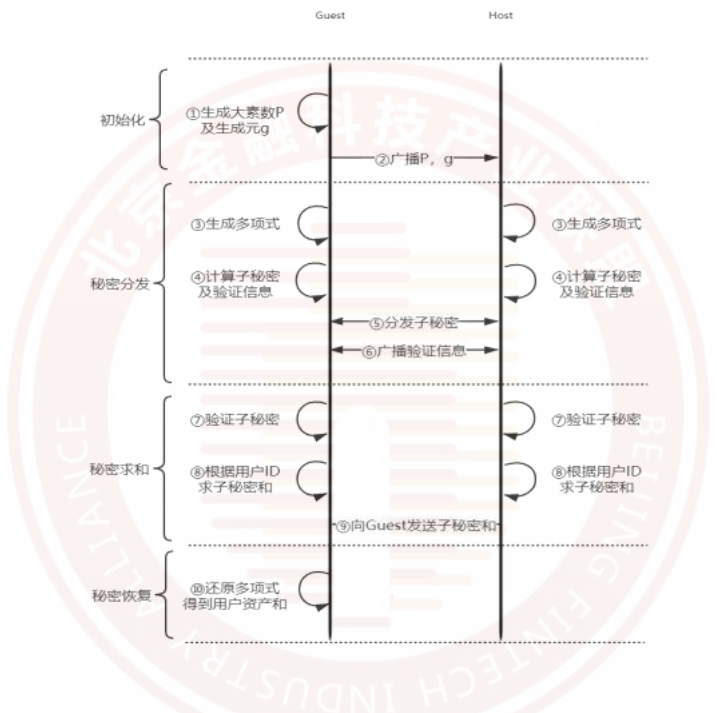


图4-21 可验证秘密共享

光大集团在集团和成员企业间搭建联邦学习平台，对在多家成员企业的交叉用户进行资产求和展开实践尝试。重点考察任务规模和开销的关系，并检查准确度。具体的实现流程分为初始化、秘密分发、秘密求和、秘密恢复四个阶段。

实验表明，随着参与方的增多，需要计算和通信的数据量都随参与方数量线性增长。计算代价与通讯代价都与总体数据传输

量成正比，扩展性良好。利用基于可验证秘密共享的安全多方隐私求和方案可以准确计算用户资产和，并且总耗时与数据传输量呈线性关系。

案例二：基于多方计算的数据交易平台

主要应用技术：多种技术融合

基于多方计算技术的数据安全融合平台可作为各行业数据的归集、处理和共享的支撑平台，在保障数据安全的同时，充分释放其计算价值。例如，日前成立的北京国际大数据交易所（简称“北数所”）应用以多方计算为核心、其他隐私计算技术及区块链存证相结合的大数据交易平台，打造了创新的数据交易模式和规则，在保障数据隐私安全的前提下实现了跨域大数据交易流通。具体数据交易模式如图 4-22 所示。

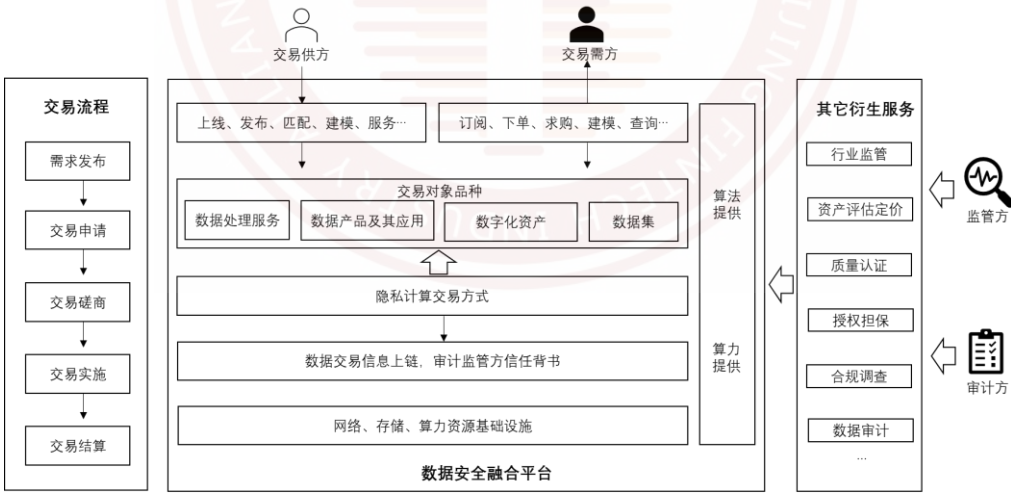


图 4-22 北数所数据交易模式

其中，交易供方为数据交易提供数据和算法，并通过交易平台发布。交易需方在北数所发布其需求并提交相关数据应用申

请，并根据自身需求在交易平台上选择相应的数据和算法，形成计算合约，最终完成交易。交易过程可能涉及第三方服务，以及为保证交易安全将交易关键信息上链等行为和操作。

在交易中，大数据交易平台的技术流程如图 4-23 所示。

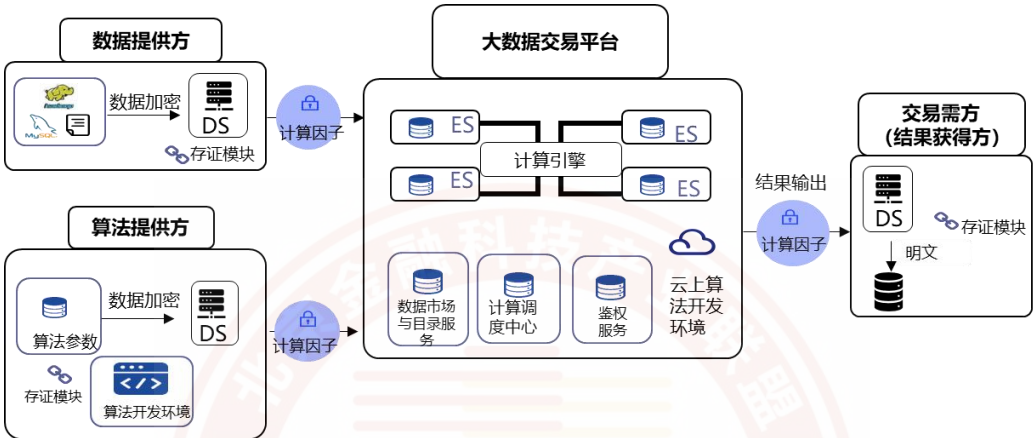


图 4-23 基于多方计算的数据交易平台技术流程架构

具体技术流程如下：

第一，数据和算法输入环节：数据提供方将交易需方需要融合的外部数据通过输入节点（提供 MPC 数据输入功能 DS）以密文形式，即“计算因子”提交至多方计算引擎；同时，算法提供方将算法参数加密提交至多方计算引擎。

第二，数据计算环节：多方计算引擎对计算因子进行预处理后分发至各计算节点（提供 MPC 数据计算功能 ES），确保每个计算节点仅能获取单个输入节点计算因子的部分“切片”。各计算节点完成各自计算后将结果提交至计算引擎，由多方计算引擎将计算结果切片后分发给若干个计算节点（每个计算节点仅能收到部分切片）进行计算。

第三，数据输出环节：结果获取方的输出节点对计算因子进行解密，获取明文融合结果。随后各计算节点销毁计算因子、计算中间结果和最终结果数据。

应用该平台，北数所首先实现了北京市政务和金融数据的交易流通，采用授权调用、联合统计及建模等方式，开展政务金融领域的协同应用服务，为政务数据向金融业开放交易、赋能金融服务做出标杆性示范。利用该平台，商业银行可安全融合政务和金融数据，在包括小微企业融资、普惠金融等诸多业务中实现对客户风险的精准刻画，在有效防范风险的同时更好地服务实体经济。

此方案对各数据交易方的价值在于：

政府方面：放心开放政务数据的流通，赋能金融行业。

银行方面：利用政务数据全方位评估企业客户信用风险，降低坏账风险，助力小微企业融资及普惠金融发展。

客户方面：破解融资难、融资贵的问题。

案例三：信贷全流程可信计算

主要应用技术：可信执行环境、区块链

某银行积极响应国家号召，积极探索小微企业信贷新模式，过去小微企业因为可抵押物少、经营不稳定，因而风险定价较高，银行因为害怕坏账追责，而“不敢贷”，一味的强求银行放贷也是违背客观经济规律的。

一方面，如果银行能深入到企业经营中去，找到可以管理和

定价的风险，使得银行小微信贷业务更加有据可依，那么银行就将可以通过金融产品创新、风险管理方式创新，甚至是开放银行能力的创新等多方面来切实支撑小微企业的发展。

另一方面，深入每个小微企业经营过程，如果没有可行的银企协作方式，将是成本高昂，不可行的，此前一些银行尝试的银税互动、发票贷、仓单贷、各类应收应付贷等，由于数据来源的不确定性，以及数据本身持续性问题，造成了银行这类产品大多只能在贷前应用部分数据进行风险分析建模，但无法实现更加高效、持续性的风险管理机制。

为该银行通过与光之数合作，利用区块链及 TEE 可信计算环境技术，在贷前贷中贷后等业务阶段安全引入了价值含量极高的数据依据。合作示意如图 4-24 所示。

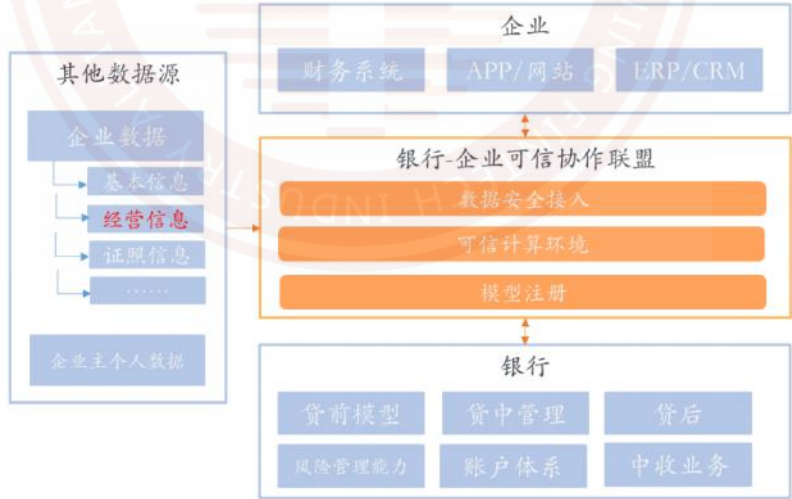


图4-24 合作示意

一是为该银行安全引入了包括国家税务（发票）、企业征信

等在内的数据。

二是利用可信计算环境，在小微企业实现源端数据安全计算，及时可信的利用企业多种经营数据进行计算分析，回传所需指标和模型计算结果。

同时，基于可信计算，银行实现了对小微企业的安全能力开放和输出，如银行账户体系、支付、票据等中收业务、风险管理能力等。从而形成了从授信者向企业经营辅助者这一综合金融服务提供商定位的转变。

案例四：多机构信息互联

主要应用技术：多方安全计算、区块链

中国人民银行南昌中心支行与趣链科技合作，打造的江西省金融业数据共享平台，平台连接中国人民银行南昌中心支行、全省金融行业大型机构和小型机构、省内相关政府部门，实现隐私保护前提下的跨机构数据共享与模型计算，实现在参与方本地数据不出库的情况下运行联邦计算模型，赋能商业银行精准的业务决策。

依托区块链与多方安全计算等前沿技术，趣链科技联合成都建信金科为中国人民银行成都分行打造了信盟链-风险信息协同共享平台（合作示意如图 4-25 所示），连接四川地区数十家商业银行及多个政府部门形成风险共治联盟，在保护各方数据隐私的前提下共享银行卡异常开户、异常交易、公安司法等风险黑名单信息，适用于防范电信诈骗、信用卡犯罪、欺诈骗保、多头借

贷等诸多业务场景，实现风险联防联控效果。

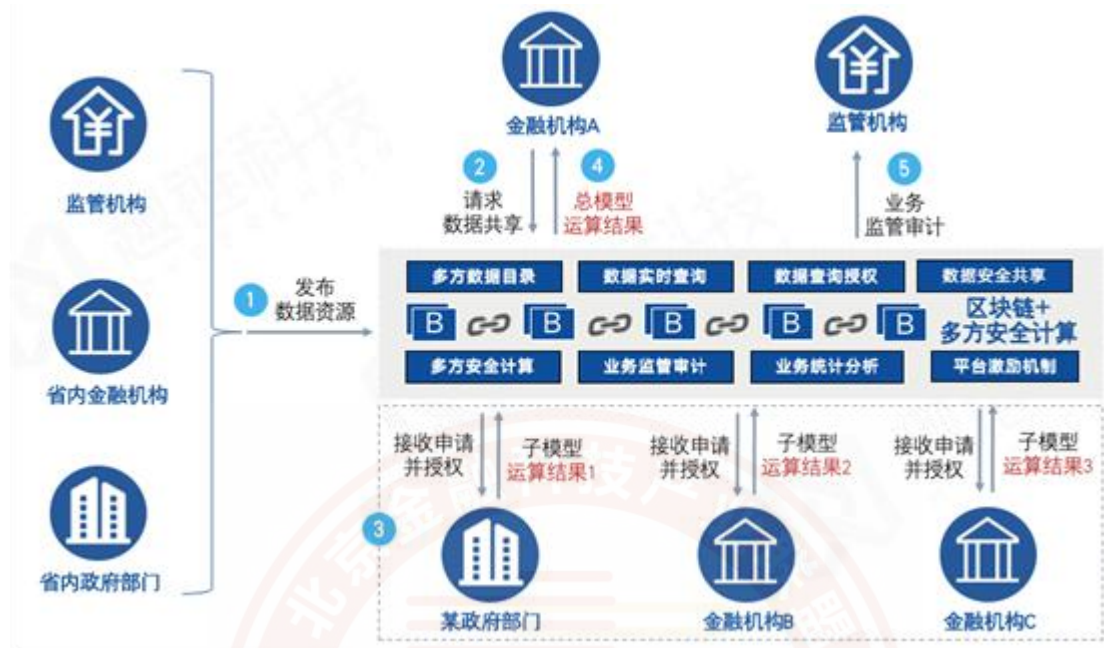


图 4-25 合作示意

案例五：全生命周期的大数据隐私计算智能风控平台

主要应用技术：多种技术融合

国内某商业银行与洞见科技合作，建立了贯穿普惠金融信贷全生命周期的大数据隐私计算智能风控平台，将该银行的行内信贷用户申请信息、存款、理财、行为偏好等数据和工商、运营商等其他合作方数据进行了安全融合，极大丰富了普惠金融信贷用户风控数据特征维度，扩大了其他合作方数据开放程度，从安全融合到智能发现，再到决策应用，全面提升了多方大数据在行内的普惠金融风控应用价值和效率。业务应用流程如图 4-26 所示。

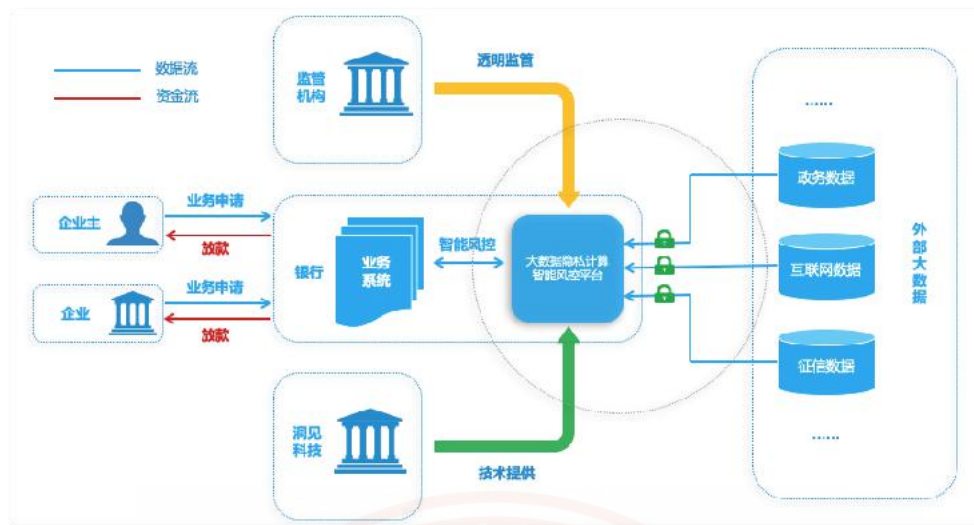


图 4-26 大数据隐私计算智能风控平台业务应用流程

通过本案例，该银行联合了包括工商、税务、水电、司法、运营商、征信机构等十余家跨行业数据源提供的上千个数据维度的外部大数据进行中小微和个人经营贷产品的风控，经测算，基于多方安全计算进行训练得到的信用评分模型无损于传统方式得到的模型，甚至 KS 值还比原来高出 2.5%，并且有效优化了银行建模分析决策路径和信贷风控流程，平均审批效率较之前提升了 30%，不仅大幅提升了行方的风险管理水平，而且也极大优化了客户申请体验。

该案例的创新是一种金融机构数字化与智能化转型的安全保护新范式，能够为银行等金融机构基于多方大数据的普惠金融业务提供更安全、更高效、更精准的安全数据智能与隐私保护计算技术赋能。通过将隐私计算、人工智能、大数据等前沿技术有效结合，搭建桥梁打通“数据孤岛”，赋能数据价值的安全释放，不仅助力银行科技基础平台能力得到了大幅提升，也极大促进了

科技创新技术未来在精准营销、反欺诈、反洗钱等其他金融业务应用场景的工程落地和实践。

案例六：可信数据共享交换

主要应用技术：可信执行环境、区块链

广东省推动深圳先行示范区数据要素市场化配置改革试点，希望实现数据要素市场化配置和高效的数据流通。但是会面临如何确定数据归属权、如何确定数据使用权责、如何保护个人隐私信息等问题。

依托于腾讯云的基于区块链和 TEE 的可信数据交换技术，深圳市龙华区建设了可信数据共享交换服务系统，实现政府机构与政府机构、政府机构与企业、政府机构与金融机构之间多个网域的安全可信数据交换及业务协同。

在功能设计方面，可信数据共享交换服务系统基于区块链分布式架构、共识记账、账本共享以及不可篡改的通道能力，连通链上及链下的数据，在可信硬件执行环境中，实现基于多方数据的协同计算。在保证数据安全、隐私保护和平等对待的前提下，系统实现了数据使用可管理、计算过程及结果可信，更好地实现业务创新。

可信数据共享交换服务系统的全功能流程，如图 4-27 所示。

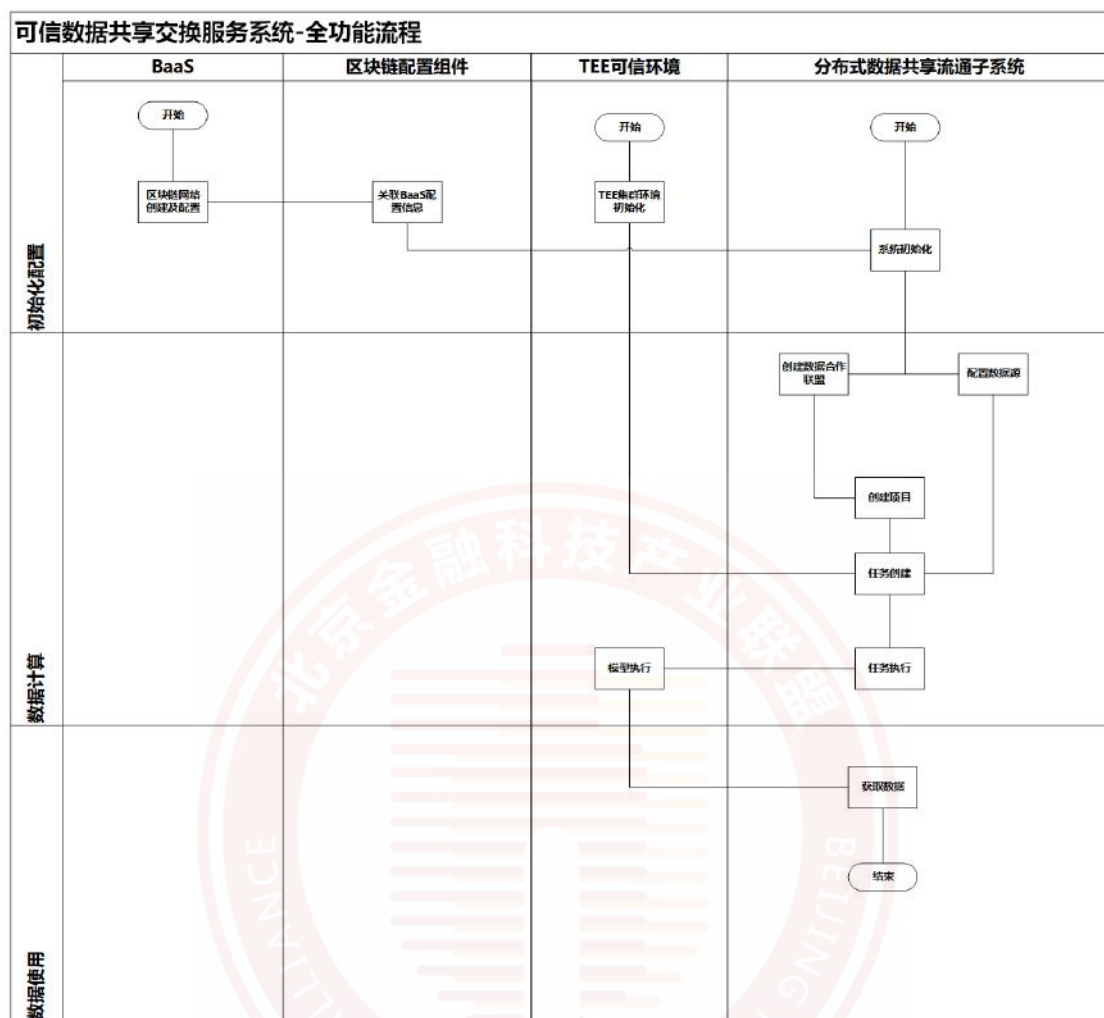


图4-27 可信数据共享交换服务系统的全功能流程图

该系统主要包括三部分的能力：区块链配置组件、TEE 可信计算框架、分布式数据共享流通子系统。第一部分是区块链配置组件,用来实现对区块链 BaaS 平台的隔离,通过配置组件这一中间层,可以实现对不同区块链底层网络的管理与配置。第二部分的 TEE 可信计算框架用来提供可信计算的基础支撑与实现能力,包括远程证明、可信信道、数据密封、密钥管理及 SDK 等。第三部分是布式数据共享流通子系统,用来实现对数据共享业务流程

的可视化操作与实现，系统主要功能如下：数据合作联盟管理、可信计算集群管理、数据管理、模型管理、项目管理、任务管理及系统设置功能。

该系统达到了如下的效果：

三权分置：数据所有方、数据使用方及算法提供方三权分置，数据所有方不再透露或传输原数据，各自加密状态下分散计算后再融合协作得到联合计算结果，实现“数据可用不可见”，在同业信任不足的情形下能获得数据合作计算的价值。

数据可用不可见：通过数据流通一方面让数据所有方既能够分享数据的价值、保护数据的安全和数据的隐私，又能让数据使用方既能得到业务所需要的数据。

生态互通：通过系统建立数据互联互通的生态、建立数据和人工智能算法互联互通的生态和数据提供方，数据使用者和算法提供方的生态。为未来数据治理，数据交易，政务数据共享提供基础设施。

通过该平台实现了深圳市龙华区的水务局水涝模型的可信计算，通过大数据训练，保证模型的精准，通过区块链与 TEE，将数据所属权与使用权分离，在安全环境下进行计算，提前预警，保障百姓生命财产安全。

（八）平台应用

案例一：金融机构隐私计算平台

主要应用技术：多种技术融合

交通银行在深入研究隐私计算领域的基础上，搭建隐私计算平台，并结合大数据、人工智能、知识图谱等技术，实现与电信运营商、支付机构等多方进行数据保护、跨界协同计算与协同建模，打造数字经济时代的新基础设施，释放数据要素红利，加快数字化转型和范式变迁^[22]。

交通银行隐私计算平台采取私有化部署方式（如图 4-28 所示），平台架构分为计算层、数据层、组件层、服务层和业务层，综合使用秘密共享、混淆电路、不经意传输、同态加密、联邦学习、隐私集合求交 (PSI)、隐私信息检索 (PIR) 等隐私计算技术，提供数据安全匹配、安全联合计算、完全联合建模、安全查询等跨机构间可信数据协作能力，释放金融数据价值，助力金融业务创新与增长。

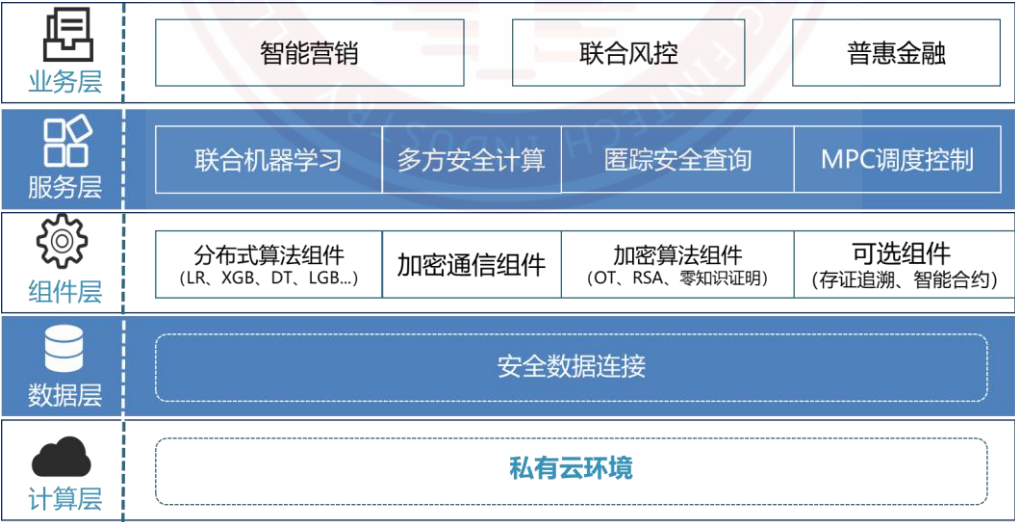


图 4-28 架构示意

平台在技术方面支持国密算法、无第三方直连网络，较好的

满足了灵活应用的要求。一是通过采用开放式隐私计算框架，遵循当前隐私计算领域行业标准及相关团体标准；二是通过支持国密算法，包括 SM2、SM3、SM4 等，有效保证了平台的性能与安全性，更好满足国内金融业务场景的合规性要求；三是平台支持弹性扩展，较好的满足了金融实际隐私计算业务扩展需求。

此外，平台通过融合图计算与隐私计算技术，实现在保护各自数据的条件下跨机构间（如银行与运营商）数据安全融合，构建联合关系图谱，打破图计算的数据边界，识别更复杂、更全面的关系链条以及欺诈风险，为金融风险防控提供线索和依据。

在操作实用便捷方面，平台提供了安全可视化模块，将复杂的安全计算原理黑盒透明化，使平台使用人员可以实时可视了解计算流程、进度，高效管理相关内容，提高安全计算技术的安全可解释性。同时，平台支持图形化建模，提供拖拽式算法流程编排等技术，降低了平台使用者的使用技术门槛，使数据分析人员可简单高效地完成多种跨机构间的安全统计、安全建模、安全预测、安全模型部署等工作，灵活应对多样化应用场景。

目前，平台已经发展出多个应用场景，并充分依托中国人民银行“监管沙箱”等开放包容的创新环境，先行先试，积极探索隐私计算技术在金融业务场景中的应用，为进一步释放数据要素红利，加速数字化转型提供实证参考。

五、问题与建议

（一）风险与挑战

1. 面临的风险

在隐私计算中，各参与方互信问题以及安全性问题格外突出。如何在缺乏互信的场景下建立各参与方安全可靠的协作机制，是实际应用中亟待解决的问题。隐私计算金融应用的主要的风险来自于以下几方面：

（1）在特定情况下，是否能保证数据安全

在多方安全计算中，不同于联邦学习或特定场景的多方安全计算的算法是固定的，通用多方安全计算的应用逻辑通常具有多样性，会根据具体的统计分析业务需求来组合底层多方安全计算的基础算子，以实现每次计算任务的应用算法逻辑。这就带来了不可控性，可能会出现底层基础算法是安全的，但组合后的应用算法逻辑会泄露某一参与方的隐私数据，例如 $A+B+C-B-C$ 。如何通过各种手段来确保应用算法逻辑安全性，是多方安全计算技术在金融行业广泛落地应用需要解决的一个重要难题。

一方面，参与方所提供的输入缺乏相应的质量验证机制，恶意的参与用户可能会提供虚假的模型参数来破坏学习过程。如果这些虚假参数未经验证便聚合到整体模型中，会直接影响整体模型的质量，甚至会导致整个联邦学习过程失败。另一方面，模型

中间参数在传输以及存储过程中的隐私性需要进一步保护加强，近期的一些研究表明恶意的用户依据联邦学习梯度参数在每一轮中的差异，通过调整其输入数据逼近真实梯度，能够推测出用户的敏感数据^{[23] [24]}。

从技术角度来看，部分通用算法（如全同态加密、零知识证明等）仍需要技术突破，才能实现更好的性能和安全性平衡。密码学算法的安全假设需适配实际应用，同时整体安全性面临量子攻击考验。

（2）较高的成本限制了应用快速落地

基于密码学的隐私计算多数依赖于分布式的计算架构，对项目实施存在较大的设计和部署方面的影响，例如在一个联合建模的场景下，需要合作方提供计算节点的本地化部署，增加了一定的成本。隐私计算的部分技术，如多方安全计算、同态加密的性能，在一些场景下依赖于硬件堆砌，也对客户增加了较大的经济和人力建设成本。此外，隐私计算的各类产品和工具依赖于底层的各类密码学和安全算法，上手难度较大，人力成本较高。这些成本的承担和削减都会给应用落地带来潜在的风险。

（3）多样化需求能否被完全满足

数据作为生产要素，受到重视的时日尚浅。数据权责仍然不明确，对技术实施增加了相当的难度；同时，不同的金融机构在不同的计算场景下，隐私需求也各不相同，并且针对特有的应用场景，需求方往往会针对隐私提出多种不同的业务要求和符合实

际业务场景的保护模式，能否满足对方案设计以及算法定制和研发提出的更高要求还有待考量。

（4）隐私计算市场还有待成熟

目前隐私计算市场的整体认知仍然较浅，金融机构对隐私计算技术的应用还在起步阶段，技术提供方需要花费极大的人力和物力成本与需求方进行技术沟通，解释成本非常高。同时，尽管各类数据隐私相关的技术标准纷纷落地，但行业仍然面临标准化进程较缓，高认可的技术规范和标准较少，以及可操作可实践的指导性标准匮乏的窘境。国内目前传统的粗放型数字经济模式（明文数据买卖）仍有延续性，数据灰色交易无法断绝。

2. 面临的挑战

数据是现代商业与个人的核心价值和重要资产。在金融行业中跨机构、跨行业的数据融合、联合分析和建模的需求日趋增加，但由于数据本身可复制，易传播，一经分享无法追踪，数据资产的确权困难，商业化被严重制约^[25]。目前，隐私计算在金融行业有诸多可落地的应用场景，但由于隐私计算技术仍在完善过程中，金融行业的隐私计算应用也面临着一些挑战。

（1）隐私计算在算力上有待提高

虽然目前隐私计算的性能已经大大提升，但由于其加密机理复杂、交互次数多，当流通的数据量较大或结构较为复杂时，会影响计算的时效性，计算效率问题仍然是重要瓶颈。特别是对于

复杂算法的模型训练效率与明文计算有较大差距，模型训练速度仍然有很大的提升空间。目前的办法有算法优化、硬件（如GPU/FPGA）加速等^{[26][27]}。

（2）隐私计算恶意模型的支持

目前大多数隐私计算产品只考虑了半诚实模型，即假设每个参与方都执行所规定的算法，他们可能试图通过从协议执行过程中获取的内容来推测他方的隐私，但不会有不遵循协议的恶意为，而一些产品并未考虑恶意模型，且即使是一些满足恶意为下安全性的产品，性能和效用都有待提高。

（3）不同平台互联互通困难

由于缺乏标准化建设指引，各金融科技企业采用的隐私计算技术各异、算法各异、数据格式和接口不统一，导致各隐私计算平台之间暂时无法实现完全互联互通，如果各参与方使用相互割裂的不同平台，则无法实现数据价值的有效融合并发挥价值。

（4）数据规范性和数据质量难以完全支撑业务

隐私计算对参与方的数据规范性和数据质量要求较高，部分企业本身科技力量较弱，自身的数据治理水平有欠缺，在引入这些外部机构数据进行联邦建模时，可能会难以发挥隐私计算技术的价值。

（5）隐私计算相关标准规范有待健全

国内隐私计算行业仍处于起步发展阶段，目前金标委发布了《多方安全计算技术金融应用规范》，但在联邦学习、可信执行

环境等方面相关的标准仍在制定中。监管部门、金融机构、金融科技企业等开展隐私计算应用时缺少相关文件指导，容易各自发展自成体系，不利于各机构开放融合以及行业健康发展。

（6）数据开放程度较低

由于认识不到隐私计算的价值，缺乏开放数据的动力，各金融机构不愿意共享开放；或者由于缺乏相关的法规文件，担心出现问题后责任认定不清，各金融机构不敢共享开放；由于行业的技术缺少相应的实践，各金融机构不愿共享开放^[28]。

（二）发展建议

为保障隐私计算产业稳定可持续发展，促进隐私计算技术在金融领域的稳妥应用，需要从政策、标准、技术发展、产业发展等方面加强建设。

1. 政策建议

为加快培育金融业的数据要素市场，加强行业数据共享流通合作、提升金融数据资源价值以及提高数据安全保护能力，提出如下建议：

（1）建议监管部门提出开展数据隐私安全共享的指导意见，在隐私数据保护的前提下开展数据合作，进而逐步推动隐私计算在行业的应用发展。

（2）建议监管部门鼓励机构开展试点探索，快速推动并形

成适合行业数据发展的路径。

（3）建议鼓励金融机构以及科技公司加强多方安全计算、联邦学习等隐私计算在数据流通等技术攻关和试点应用。

（4）建议建立行业关键基础设施，促进行业数据流通，以及对接跨行业数据，同时增强数据安全预警和溯源能力。

2. 标准化建议

为加强隐私计算在金融业的合规应用，需要从技术要求、评测要求以及实施指南等方面进行标准化建设。

（1）技术要求类标准。隐私计算技术要求类标准，建议包括应用技术规范、互联互通规范、技术安全规范、性能要求规范等。

（2）测评方法类标准。隐私计算测评方法类标准与技术要求类标准配套，指导隐私计算技术的测评认证，建议包括应用技术测试方法、互联互通测试方法、技术安全测试方法、性能测试方法等。

（3）实施指南类标准。针对隐私计算技术在金融领域内的应用，建议制定实施指南类标准。如编制应用实施指南，指导隐私计算技术在智能风控、智能营销、反欺诈等场景中的应用。

3. 技术发展建议

隐私计算兼具理论研究和实际应用价值，为大数据安全和隐

私保护计算提供了一条重要的技术路径。面对隐私计算需求快速增长的大环境，有必要尽快推动隐私计算技术发展。

（1）加快密码算法和安全硬件验证。一是加快支持国家密码管理部门认证的密码算法，并推广应用。二是基于安全硬件的 TEE 方案是允许应用程序实现一个被称为 Enclave 的容器，在应用程序的地址空间中划分出一块被保护的区域，为容器内的代码和数据提供机密性和完整性的保护，免受拥有特殊权限的恶意软件的破坏^[29]。

（2）提升基础性能，落地更多复杂算法。最近一年，各个厂商从算法设计、工程优化等方向大幅提升了隐私计算技术的基础性能。然而现有的计算产品大多仍只支持逻辑回归、XGBoost、线性回归等常见机器学习算法以及少数简单的 AI 算法，海量数据模型训练及更多复杂 AI 算法的联邦化等还需优化和突破。随着软硬件技术不断突破和算法不断优化，隐私计算门槛会越来越低，支持的复杂运算任务将会越来越多，易用、通用的隐私计算产品将会是大势所趋。

（3）提升安全协议的扩展性和强度。现有研究中参与方比较少，如何设计算法支持成百上千个参与方的大规模计算、如何设计协议支持复杂网络模型下的计算，是提高隐私计算技术实用价值的关键。学界、业界等各方在积极研究通信效率更高、计算速度更快的安全协议，全同态加密、可信执行环境、使用 GPU、FPGA 加速都是有益的尝试，也是未来重要研究方向。值得注意

的是，谷歌的 Craig Gidney 和瑞典斯德哥尔摩 KTH 皇家理工学院的 Martin Ekerå 一项最新研究成果表明，2048 位 RSA 密码使用量子计算机能在八个小时被暴力破解，而这种密码用超级计算机，破解需要 80 年。计算机硬件的提升为安全协议效率的提升创造了可能性，同时也给安全协议的强度提出了更高要求。

（4）加快隐私计算相关技术融合发展。区块链可作为隐私计算技术的监控和认证保护，区块链技术的去中心化、不可篡改等技术特性，将进一步增强隐私计算的安全存储和隐私保护性。而以可信执行环境为主的可信计算技术一般作为整个隐私计算产品中的硬件扩展，在全密文计算等技术的基础上进一步加强运行环境安全，从技术上保障安全模型所需条件得到满足。现有的隐私计算产品在可信计算上落地稍多一点，但是在去中心化场景如用区块链存证实际落地的业务场景还有待进一步发展。

4. 产业发展建议

隐私计算技术正处于快速发展过程中，关键技术已经取得许多突破，但是产品化方面还较为欠缺。在当前数据要素流通的诉求下，对隐私计算技术提出了较高的要求。为了能够支撑金融业数据创新发展，支撑普惠金融服务，应采取有效措施推动产业发展。

（1）建议加强产学研用联动，充分吸收科研成果，迅速转化为实践应用，在实践中指导技术发展方向，推动隐私计算技术

产业快速迭代。

（2）结合国家对数据安全监管的政策要求，形成相应的技术解决方案，应对数据隐私安全应用要求，确保个人隐私信息安全，符合国家法律法规要求。对于企业应用隐私计算来说，宜根据企业自身情况，建立合理的内部合规制度，完善管理流程，做好数据安全内控。

（3）加强与人工智能、区块链、边缘计算等金融科技进行融合发展，基于云计算平台形成科技服务能力，形成体系化的技术解决方案。

（4）摸索和探索隐私计算应用商业模式，联合数据使用方、数据提供方等多方参与联邦建模，逐步建立数据服务生态，探索创新发展商业模式。

参考文献

- [1]. 分布式capital:《分布式课堂: PlatON创新研究院夏伏彪解析隐私计算技术》, 2020-11-16。
- [2]. 张宁:《基于流量预处理的新型流关联技术的研究》, 2017-05-18。
- [3]. 陈志德:《不经意传输协议》, 2005-06-30。
- [4]. 崔泓睿:《多方安全计算热点: 隐私保护集合求交技术 (PSI) 分析研究报告》, 2019-08-14。
- [5]. 崔泓睿:《带隐私保护的集合交集计算协议的发展现状综述》, 2019-03-31。
- [6]. 周俊:《联邦学习安全与隐私保护研究综述》, 2020-02-15。
- [7]. 杨玉龙:《基于同态加密的防止SQL 注入攻击解决方案》, 2013-12-09。
- [8]. 信通院:《隐私计算白皮书(2021)》, 2021-7-20。
- [9]. 刘华东:《处理个人信息应事先告知并取得同意》,《光明日报》2020年10月14日, 第8版。
- [10]. 洞见科技:《从《个人信息保护法》看隐私计算的科技向善应用——关于隐私计算应用场景的合规路径浅析》, 财经网, 2021年11月1日。
- [11]. 腾讯研究院:《腾讯隐私计算白皮书2021》, 2021年4月19日。
- [12]. 移动支付网:《央行发布<金融业数据能力建设指引>明确5大基本原则》, 2021-03-08。
- [13]. 科技之佳:《金融科技热点应用突破与创新白皮书》, 2020年11月8日。
- [14]. 李敏:《我国银行卡清算市场的发展现状研究》,《现代商业》2016年第5期。
- [15]. 张艳艳:《“联邦学习”及其在金融领域的应用分析》, 2021年1月15日。
- [16]. 陈菁菁, 谢烨烨, 蒋舟阳, 黄阳:《共生理论视角下浙江小微企业信贷约束机理》,《时代经贸》2016年12期。
- [17]. 何竞平:《互联网广告管理新规的传播学分析——基于德弗勒互动过程模式的探讨》,《青年记者》2018年3月下。
- [18]. 连一席:《区块链研究报告: 从信任机器到产业浪潮还有多远》,《发展研究》2018年第8期。
- [19]. 陈垣桥:《全球银行业反洗钱监管现状与趋势分析——基于2017.1—2020.6 全球银行业反洗钱监管处罚数据》, 2020-12-14。

- [20]. 《反洗钱任重道远：行业平均误报率仍有95%，AI算法还在探索期》，《21世纪经济报道》，2020-3-25。
- [21]. 石源，张焕国，赵波，于钊：《基于SGX的虚拟机动态迁移安全增强方法》，《通信学报》，2017年第38卷第9期。
- [22]. 谭培强，谢谨：《多方安全计算金融行业应用初探》，《金融电子化》，2020年第12期11-12。
- [23]. 云悦科技：《人工智能特别是联邦学习用于提高区块链技术的安全性之探讨》，2020-03-17。
- [24]. 崔泓睿：《带隐私保护的集合交集计算协议的发展现状综述》，2019-03-12。
- [25]. 头条：《ARPAx京东数科：隐私计算如何赋能未来金融数据共享》，2020-05-26。
- [26]. 晨山：《隐私泄露、数据孤岛、数据滥用……「隐私计算」如何解决大数据行业共性挑战？》，2021-01-04。
- [27]. 云悦科技：《安全多方计算的起源与基础介绍》，2020-02-09。
- [28]. 闫桂勋：《数据共享安全框架研究》，2019-05-04。
- [29]. 陈琨，王国赛，李艺：《多方计算在个人征信领域中的应用》，《数字经济与社会》，2021年6月8日。

附录：隐私计算技术平台产品

1. 数牍平台产品—基于隐私计算的多方安全数据协作平台

数牍科技开发的基于隐私计算的多方安全数据协作平台，包括底层加密协议、系统架构及安全、算法功能、场景应用等四层，如图 1 所示。



图 1 系统架构及功能分层

底层加密协议使用各种国际通用和国家密码局发布的加密算法，覆盖同态加密、秘密共享、不经意传输等技术。架构部分具有高可用、高扩张特性，提供安全的任务调度、调用、传输等功能。协调多个参与方（企业，政府机关及其他组织）进行 ID 融合、AI 建模、统计分析等具体功能。平台提供全流程的数据科学解决方案，包括数据打通、联邦分析、联邦建模、联邦预测

等功能，可以有力支持实际落地业务场景，例如用户画像，安全查询，ID 打通，联合征信等。

2. 瑞莱智慧平台产品—RealSecure (简称 RSC)

北京瑞莱智慧科技有限公司推出的隐私保护计算平台 RealSecure (简称 RSC) 是基于多方安全计算、联邦学习、匿踪查询等技术打造的数据安全共享基础设施。针对有隐私保护和联合建模需求的企业级用户，通过打通数据孤岛，将计算环节移动到数据端，实现数据可用不可见，解决多家机构数据合作过程中的数据安全风险和隐私泄露问题。平台保证在数据不出库、无中间方作为协调者的前提下，完成多方参与的模型训练过程并获得可投入生产的模型，平台保证模型效果与将数据汇总到本地建模效果一致。

RSC 由数据生态、平台系统、业务方案三大核心板块组成：其中数据生态包含支付、运营商、互联网、政务等多类型多场景数据，形成能够覆盖金融机构各类业务场景需求的安全计算数据生态体系，保障金融机构在需要数据支撑业务建模时通过 RSC 平台实现无缝、安全、合规对接，拓宽自有数据可利用维度。平台系统包含联合建模、联合统计和匿踪查询三个功能模块，自主研发的全同态加密、联邦 AI 编译器等技术使平台具有极高的性能、易用性和可拓展性。业务方案即针对智能风控、智能营销、流动性管理等场景提供成熟的业务解决方案，金融机构可在需要解决

对应的业务问题时通过 RSC 平台实现快速落地应用，切实地解决业务问题。

3. 洞见科技平台产品—洞见数智联邦平台（INSIGHTONE）

如图 2 所示，洞见数智联邦平台（INSIGHTONE）是洞见科技基于隐私计算和区块链技术独立自主研发的金融级隐私保护计算平台产品，其目标是构建数据智能联邦，让数据融合价值更安全地释放。名字中的 INSIGHT 承载了对于数据产生智能的洞见，ONE 代表着多方数据价值在智能层面安全达成融合统一。目前，INSIGHTONE 平台是国内少数通过国家工信部中国信通院“多方安全计算、联邦学习、隐私计算性能与安全”三项评测认证的产品，已通过国家公安部信息安全三级等级保护测评并获得国家网信办区块链信息服务备案，还荣获“2020 安全数据智能最佳产品”奖、“2020 最佳隐私计算平台”奖等多项荣誉。



图2 系统架构及功能分层

洞见数智联邦平台（INSIGHTONE）是涵盖了数据要素管理、计算引擎管理、应用服务管理及可信网关“三横一纵”技术架构的一体化科技服务平台，具备“高安全、高性能、高兼容”三大优势特性：一是自研无第三方联邦学习（NTP-FL）技术，解决多方联合建模中的第三方可信风险问题；二是自研快速联邦学习技术，相比开源算法有数十倍的速度提升，解决联合建模中的性能问题；三是融合“MPC+FL”双计算引擎，适配不同计算场景，并在一定标准内支持异构计算框架之间的互联互通，如图 3 所示。



图 3 功能模块示意

基于洞见数智联邦平台（INSIGHTONE），洞见科技已打造“Insight+”系列硬核技术产品体系，包括：InsightMPC、InsightFL、InsightML、InsightPIR、InsightChain、

InsightStrategy、InsightGraph 等，通过立体组件化数据智能产品矩阵，以及匿踪查询、集合运算、联合统计与智能建模等应用服务矩阵的构建，全方位满足各行业数据安全融合与隐私保护计算的需求，已在金融、保险、政务等业务场景取得了数十个落地案例，为客户的数据安全流通赋能，实现数据的“可用不可见”。

4. 富数科技平台产品—Avatar 安全计算平台

Avatar（阿凡达）安全计算平台是上海富数科技有限公司自主研发的一站式企业级安全计算平台，集成多方安全计算、联邦学习、匿踪查询等核心技术，提供企业级的数据安全匹配，安全联合计算、安全联合建模、匿踪查询等跨机构间可信数据协作能力，释放数据价值，助力业务创新与增长。Avatar 安全计算平台采用开放式框架，有效实现算法模型与平台的解耦，支持热插拔安全算子，算子能够以类似乐高模式进行模块组合，灵活搭建，大幅度提高联邦学习和多方计算流程的灵活性，解决应用场景多样性、建模人员建模习惯差异化问题，更好满足实际业务场景的需求。Avatar 平台的架构如图 4 所示，主要包括安全计算引擎模块、安全计算接口模块、数据资源模块、管理运营模块以及多方互联互通模块。

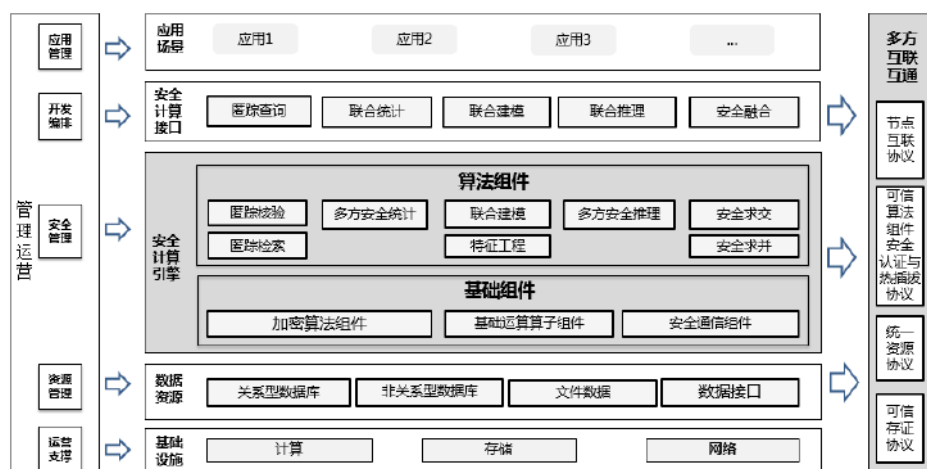


图4 系统架构及功能分层

Avatar 安全计算平台基于自研的安全算法，同时支持有可信第三方和无可信第三方直联组网的多种场景，并首创安全驾驶舱，让复杂的多方安全技术和联邦学习技术黑盒透明化，让安全可视化，提升安全可解释性，让用户掌握更强的系统运营能力。在此基础上，用户可以结合实际场景以及自身实际需求，通过增加相关模块（包括定制联盟区块链模块、安全策略模块、AI 计算模块等）对 Avatar 标准平台进行补充实现更多功能。

5. 蓝象智联平台产品—GAIA Cube 保护数据隐私的一站式联邦学习平台

蓝象智联专注于提供金融级隐私计算技术及产品。团队核心成员来自阿里巴巴和蚂蚁金服，深耕隐私计算联邦学习领域 4 年多，拥有极深的行业经验积累和资源优势。

公司自研核心 GAIA 产品系列，基于多方安全计算和联邦学

习技术，在 AI、密码学、大数据等技术基础上，结合丰富的行业经验，为金融、运营商、政府、互联网平台等行业提供自主研发的隐私计算产品和解决方案，实现跨行业数据建模、可信数据资产交换及分布式智能应用需求，满足行业营销、风控、大数据交易所等业务领域的需求。其中，自研的多方安全计算产品 GAIA EDGE 和一站式联邦学习产品 GAIA CUBE 更是国内首个金融级隐私计算产品。

GAIA 产品系列具备业内最快性能的 PSI 隐匿求交技术，10 分钟内可以完成 5 亿用户数据表的求交。其多算子联邦算法库可实现时间最快、精度最高、带宽最少的效果。产品内的双安全等级部署支持保护数据也保护模型的模型部署，在金融场景上线中保护客户的多的模型资产。产品支持双引擎驱动，包括联邦 AI 引擎和多方安全计算引擎，具备全栈的技术可拓展性。GAIA 系列更是可视化建模体验最优的产品平台，可以帮助业务团队完全实现零代码、所见即所得的建模体验。

蓝象智联是《多方安全计算金融应用技术规范》（JR/T 0196-2020）行业标准起草单位，同时也是工信部的多方安全计算标准的起草单位。

蓝象智联目前也正在与电信、移动、联通三大运营商形成合作，除了基于蓝象隐私计算系列产品构建运营商与其他行业（特别是金融机构）的隐私数据交互桥梁之外，蓝象智联的算法能力和数据运营能力也正在帮助运营商加大数据对外合作力度，基于

金融机构的场景需求来安全、有序地开放更多数据，帮助金融机构进行业务创新。GAIA 产品数据服务、系统架构及功能分层如图 5、图 6 所示。



图 5 GAIA 产品数据服务



图 6 系统架构及功能分层

6. 百度平台产品—百度金融联合建模产品

百度金融联合建模产品(系统架构及功能分层如图7所示),通过软硬件结合的方式,提供从数据分析、模型训练、评估到预测部署的全流程服务,拥有可信计算环境(TEE)、联邦计算(SMPC)、安全隔离域等安全计算引擎技术,同时支持私有化、公有云以及联邦三种部署模式,可以灵活匹配不同业务场景的需求。

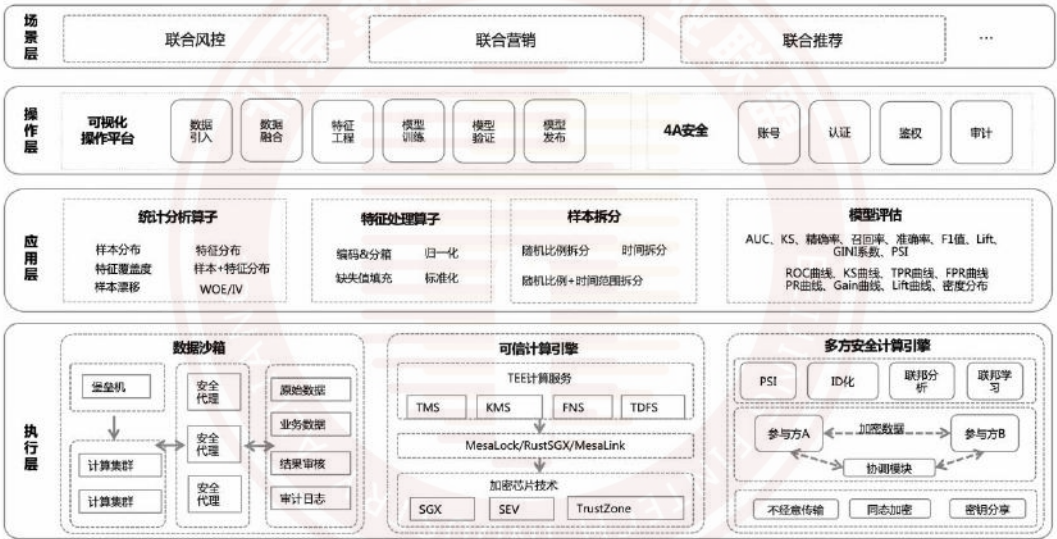


图7 系统架构及功能分层

与其他同类型产品相比,百度金融联合建模产品拥有以下优势:

(1) 全栈式安全计算引擎

产品综合运用了 TEE 可信计算、MPC 联邦学习计算及安全隔离域三种技术,拥有全栈解决方案。

（2）高性能高安全的计算引擎

在可信计算引擎方面，产品采用 MesaTEE 方案，和传统 MPC 方案相比效率更高，不会随着参与方增加导致计算性能指数型下降。

在联邦计算引擎方面，产品实现原始数据不出本地以及明密文结合计算的方式，进一步保证了安全和效率。

（3）实施方案适用范围广

产品针对不同客户设计不同的实施方案。以两方数据融合为例，企业可灵活选用私有化、联邦和云端三种部署方式，可在平台进行一系列建模操作，在模型可用并正式发布后，还可以实时调用入模数据支持模型产出最终结果。

（4）权威机构技术认证

产品通过了信通院基于多方安全计算的数据流通产品基础能力专项评测、可信数据服务基础能力专项评测等两项专业认证。

凭借以上优势，百度金融联合建模产品已成功服务多家保险公司、银行等金融机构，其主要应用场景如下：

（1）风险识别

产品提供可安全融合多方数据价值的分析与挖掘环境，帮助企业有效识别潜在风险（如在信用卡、贷款、在线支付等业务场景中的违约、欺诈风险），提高风险识别能力，减少企业损失。

（2）存量客户经营

产品提供可安全融合多方数据价值的分析与挖掘环境，帮助企业对存量客户进行挖掘与经营，优化企业营销策略，用更少的资源赢得更多客户。

（3）精算定价

金融企业往往缺乏其客户数据全貌，对客户的价格偏好缺乏全面认识，金融联合建模产品可以帮助企业整合更全面的数据，帮助企业进行更有效的精算定价。

百度金融联合建模产品凭借其独特的优势在多家金融机构的成功落地，连续两年入选《中国金融科技发展概览》典型应用案例，2020 年分别荣获 IDC 中国金融机构场景创新应用奖，亚洲银行家中国最佳数据安全技术奖、中关村金融科技论坛-2020 中国金融科技技术创新奖。并在 2020 年信通院组织的“星河”行业大数据案例征集活动中被评为“隐私计算标杆案例”。

7. 蚂蚁集团平台产品—隐私计算平台

蚂蚁集团高度重视用户隐私保护。基于此，“蚂蚁隐私计算平台”（系统架构及功能分层如图 8 所示）深耕数载，整合 AI、数据科学、密码学、安全保护技术等多技术领域，实现了在多合作方数据不可见情况下的数据安全共享，让数据不动价值动。



图8 系统架构及功能分层

技术实现上，平台满足工业级高性能隐私计算的诉求；确保在满足用户隐私保护和数据监管合规前提下，实现数据联合计算。其技术核心能力体现在：

安全性，可抵抗从半诚实攻击到恶意攻击的多种攻击模型，支持从可证安全到统计安全的多种安全等级。

计算精度，通过多种创新技术去提升密态域的计算精度，降低密态计算下的累积误差，使得计算精度和传统集中式计算的精度接近。

性能，蚂蚁结合算法创新和工程优化，可满足多种工业场景的性能需求，如在可证安全等级下的模型训练，可实现千万样本，十万级特征的运算性能。

蚂蚁隐私计算平台覆盖普惠金融、账号安全、医疗保险、疾病诊断等众多与民生息息相关的业务场景中。促进数据有序流通，实现跨域价值融合创新。

智能风控上，为金融机构提供联合多方数据源构建风控模型技术支持，有效提升风控效果。智能营销上，基于用户授权条件下助力保险公司制定、优化权益策略。智能医学诊断中，助力科研人员、医疗机构实现整合医学中心临床数据，提升医学研究的准确性、适用性和泛化能力，加强各医疗相关机构间协同研究能力。

8. 融数联智平台产品——“善数”平台

北京融数联智科技有限公司提供集“软件产品+硬件加速+专业服务”为一体的隐私计算解决方案，公司自主知识产权的“善数”隐私计算平台具备安全求交、安全求和、匿踪查询、安全建模等系列产品（系统架构及功能分层如图9所示）。“善数”平台提供了端到端的全流程可视化操作环境，让用户无需编写代码即可如顶尖专家一样深入进行各项隐私计算工作，同时我们还致力于从软件算法优化、硬件芯片/板卡以及网络等多个层面持续进行隐私计算的效率和效果提升，并且系统具有高扩展和高开放性，可以和开源的第三方系统实现对接和算子互通。



图9 系统架构及功能分层

融数联智“善数”平台已经通过了工信部信通院的评测，获得了国家软件质量监督认证中心的认证证书，产品已经服务了近百家公司，覆盖了金融、医疗、政府、零售等多个领域，得到了客户的一致好评。

9. 微众银行平台产品—FATE 企业版

微众银行从理论研究、核心算法、行业应用等多维度推动联邦学习行业生态建设。2018 年，杨强教授领衔的微众银行 AI 团队提出纵向联邦学习的方向，以面向 B 端为特征，国内联邦学习正式拉开产业化大幕。2021 年 7 月，联邦学习被首次纳入 Gartner “隐私计算的技术成熟度曲线-2021”。每年一度发布的 Gartner 曲线被视为了解全球科技新动向最具参考价值的报告之一，标志着联邦学习产业化初见成效，进入大规模应用阶段。FATE 平台是微众银行自研开源的全球首个工业级联邦学习框架，其主要目

标就是有效帮助多个机构在符合数据安全和政府法规前提下,进行数据使用和联合建模。该产品对机器学习、深度学习、迁移学习算法提供强有力的安全计算支持。安全底层支持同态加密、秘密共享、哈希散列等多种多方安全计算机制,算法层支持多方安全计算模式下的逻辑回归、Boosting、联邦迁移学习等。

微众银行 FATE 企业版为企业客户提供了符合隐私计算要求的联邦学习平台,通过隐私数据计算全流程框架,打造企业对业务数据的统一规划和管理、对合作伙伴数据的合规使用,为数据价值的体现搭建了基础设施。该产品目前已经完成国家金融科技评测中心多方安全计算金融应用评测和中国信息通信研究院多方安全计算和联邦学习产品评测。

微众银行 FATE 企业版包括了在裸金属平台、虚拟化平台以及容器平台的部署,包括 FATE 框架中常见的回归、树模型、神经网络等机器学习算法,支持横线联邦、纵向联邦和联邦迁移学习场景。除了标准的命令行工具外,企业版还提供了可视化界面,包括了可视化建模;数据、任务、模型的管理;日志和审计服务;项目管理;身份验证和授权等组件。

FATE 企业版目前已可以实现公有云和私有云的全生态应用,涵盖了金融、医疗、智慧城市等领域,并逐步向更广范围延展。系统架构及功能分层如图 10 所示。



图10 系统架构及功能分层

10. 同盾平台产品—智邦平台 (Ibond)

智邦平台是同盾科技旗下人工智能研究院基于知识联邦理论体系打造的工业级应用产品，旨在构建数据安全的人工智能生态系统。知识联邦是同盾提出的复合安全架构体系，根据联邦发生的阶段可以分为信息层联邦、模型层联邦、认知层联邦和知识层联邦。像多方安全计算是信息层联邦的典型方案，联邦学习称为模型层的联邦。智邦平台最终使命是要构建一个开放的生态，主要体现在三个层面。一是开放的任务联盟，基于智邦平台，每个机构可以参与多个任务联盟，在不同联盟中也可以开放不同的数据。二是开放的模型设计，参与模型设计的人员可以开放的加入到智邦平台，并在不同的任务联盟中针对任务需要设计模型。三是开放的模型使用，每个任务对应的模型性能效果是对外开放

的，可供使用者查询。模型使用者可以根据业务需求选择合适的模型，也可以将不同任务场景下的模型连通起来形成业务闭环。

智邦平台可通过安全的数据交换实现知识共创和共享，是打破部门数据割裂，同时确保数据安全和隐私保护的关键，在金融、保险、政务和医疗行业有很大应用潜力，也是实现智慧金融、智慧政务和智慧医疗的基础。目前的应用场景有智慧金融、智慧政务、智慧医疗、智慧城市。

通过智邦平台不仅可以打破参与方的数据壁垒充分利用各参与方的数据，同时又可以保证数据不离开参与方来保护数据隐私。智邦平台在数据隐私保护方面符合国家监管要求，也达到了国际上对数据保护的规范要求。同时研究院积极参与行业标准制定，产品目前参与了多项测评。

智邦平台包含多方安全计算、联邦学习、匿踪查询模块，支持同态加密、一次一密等加密技术，并对数据进行三级安全域划分，实现数据的可用不可见。智邦平台系统架构及功能分层如图 11 所示。



图11 系统架构及功能分层

11. 光之树平台产品—隐私计算平台

光之树科技自主研发的隐私计算平台是基于联邦学习、TEE 可信计算、多方安全计算、隐匿查询等技术打造的企业级安全计算平台，对安全和隐私有强诉求的场景提供了通用的安全计算环境 and 能力，保证计算过程中数据的机密性、完整性、可用性，从技术上打破数据孤岛，助力企业挖掘数据价值。

安全自动化建模系统内置丰富的机器学习、深度学习算法，提供模型评估报告、关键特征，满足各类应用场景需求；综合应用秘密分享、同态加密等多种密码学协议的自研架构，相比开源算法性能提升 5 倍；TEE 机密计算引擎拥有 Intel SGX、华为 Trustzone、海光 CSV 等多种 TEE 技术方案，从 OS 层优化到应用

该产品已通过已通过信通院《基于多方安全计算技术能力专项评测》和银行卡检测中心《多方安全计算金融应用技术规范评测》的测评，符合金融应用安全性要求。

(1) 产品主要功能

具体产品功能如图 13 所示。



图 13 隐私计算服务系统产品功能

隐私计算服务系统以面向上层业务需求提供全方位的隐私计算能力支撑为核心目标，通过平台化的方式，为机构搭建隐私计算应用开发能力，并能灵活调用系统隐私计算算法层和基础层的相关能力。

隐私计算算法层通过各类密码学算法(秘密分享、混淆电路、不经意传输、同态加密、零知识证明等)的工程化实现，可满足联合查询、联合统计分析、联合特征工程、联合机器学习与联合预测等多方的隐私计算模式，灵活满足不同的业务需求。

基础层可与区块链网络对接，实现对数据的使用确权、计算

过程留痕，做到算法应用、模型训练过程及结果数据可追溯。

(2) 产品实施架构

通过本地隐私计算系统与隐私协作调度平台相结合的模式，面向不同参与方按需提供软硬件相结合的多样化产品能力输出，及一体化数据安全与隐私协作解决方案。

(a) “隐私计算系统”由参与方在各自本地部署，通过隐私计算引擎配置各类隐私计算算子，灵活实现对本地原始明文数据的密态处理，确保在数据计算、协作过程中原始数据不离开本地保护，系统部署方式如图 14 所示。

(b) 通过隐私协作调度平台实现与各方本地“隐私计算系统”的对接和服务发现，更为灵活、便捷的满足不同参与方间的两方或多方数据共享和融合计算需求。



图 14 矩阵元隐私计算服务系统部署方式

13. 平安集团平台产品—蜂巢联邦智能隐私计算平台

蜂巢联邦智能隐私计算平台是数据隐私保护的一站式解决方案，为金融业务提供全方位数据安全合作服务。针对数据使用与数据流转中存在的数据共享难、数据质量可信度低、数据和模型融合难度大等问题，蜂巢平台研发基于隐私安全保护的联邦学习技术，有效帮助政府与企业机构在符合法律法规的前提下，进行数据应用和联合建模。蜂巢联邦智能隐私计算平台目前已获得软件著作权证明，被国家工信部授予“2020 年网络安全技术应用试点示范项目”，同时获得国内外多个行业荣誉，包括“2020 年 IDC 数字化专项金融奖”、“2020 年 CCF 科技进步奖”、“2020 年 BAI 全球创新奖-监管科技创新奖”，同时也是 BAI 当届获奖项目中唯一来自中国的科技项目。项目团队在国际顶级会议上发布隐私计算相关论文 10 余篇，申请技术专利 50 余项。合作示意如图 15 所示。



图15 合作示意

蜂巢联邦智能隐私计算平台包含数据生态、联邦计算、联邦建模、推理应用、平台管理、监管审计等六大模块组成，提供中心化联邦及去中心化联邦两种方案，覆盖各个业务场景中从数据管理到生产应用的全流程体系，支持同态加密、秘密分享、国密等多种加密方式。系统配备的可视化客户端与监管管理平台，具备“易用性”、“安全性”、“高效性”。蜂巢平台当前已广泛应用于跨机构数据合作、金融风控、交叉营销等场景，有效降低模型训练成本，提升训练效率。首创的联邦图谱技术，可以在不泄露双方图数据和特征信息的前提下，构建合作方脱敏的隐私图，大幅提升模型效果，完善业务策略。同时，蜂巢平台达成多个跨异构平台互联互通建模案例，在信贷风险管理场景中完成落地应用。

14. 腾讯云平台产品—云安全隐私计算（Cloud Security Privacy Computing, CSPC）

腾讯云安全隐私计算（Cloud Security Privacy Computing, CSPC）是腾讯云推出的以联邦学习（FL）、安全多方计算（MPC）、可信执行环境（TEE）等隐私数据保护技术为基础的隐私计算平台，产品针对机器学习算法进行定制化的隐私保护改造，保证原始数据不出本地即可完成联合建模，同时支持安全多方 PSI（隐私保护集合求交技术）、安全隐私查询、安全统计分析，提供基于硬件的 TEE 可信执行环境。

云安全隐私计算的特性：

（1）剥离中间方规避合谋风险

云安全隐私计算平台针对性改造主流联邦协议、剥离可信中间方，保证传递任何加密中间参数都在合作方之间直接完成，排除任何中间方包括云安全隐私计算平台本身，最大化保障合作双方数据安全。

（2）多方联邦高效稳定

云安全隐私计算在稳定的两方联邦基础上进行定制化改造，目前已支持稳定的 3 方、4 方联邦，充分发挥多方数据融合价值。

（3）单向网络策略

云安全隐私计算针对性地实现了单向网络策略，在私有化部署的基础上金融机构可以访问到数据合作方，但数据方无法访问到金融机构，这样最大化保障金融机构环境安全。

（4）产品控制台简单易用

从联邦节点注册、数据注册、项目创建、任务设置、安全求交、特征工程到算法调试等关键环节，不需要编写任何脚本，只需基于控制台简单操作即可完成，同时还可以方便地进行模型效果对比及参数调整，让联合建模成本降到最低。

（5）性能优异可扩展

基于 K8S 提供整套平台服务，基于 Apache Pulsar 提供消息通信服务，保证数据量级与通信资源的效率，同时，基于 Spark 集群提供弹性的计算服务，性能优异且可弹性扩展。

（6）轻松处理海量数据

基于腾讯云 Angel PowerFL 框架，可实现异步并发计算，轻松处理千亿级数据量。一个小时左右可以完成千万量级数据的 XGBoost 模型训练，十分钟左右可以完成千万量级数据的推理。

15. 华控清交平台产品—多方计算平台

华控清交多方计算平台通过吸收、转化清华大学的科研成果而来，采用了多方安全计算等多种隐私计算技术。主要技术特色有：

（1）易用性高。采用业界通用的 Python 和 SQL 开发接口，封装了 500 多个常用的函数库（含 Numpy、Pytorch），并提供配套 IDE，让不懂密码学的普通程序员能够直接调用密文计算函数进行编程，就像在明文环境里一样开发隐私计算应用。

（2）高性能。在密码学基础理论、底层协议、分布式计算、系统、编译、算法和芯片等方面对产品性能进行持续优化和创新。

（3）可扩展。支持数据类型、算法类型、参与方数量、任务数量等动态扩展。

（4）自主可控。从底层的基础运算（加法、乘法、比较等）开始进行工程实现，完全自主可控，具有原创性和自主知识产权。

华控清交多方计算平台允许多角色接入（平台应用示意图如图 16 所示）。平台的通用性使得这些角色可以根据不同应用场景进行调整。

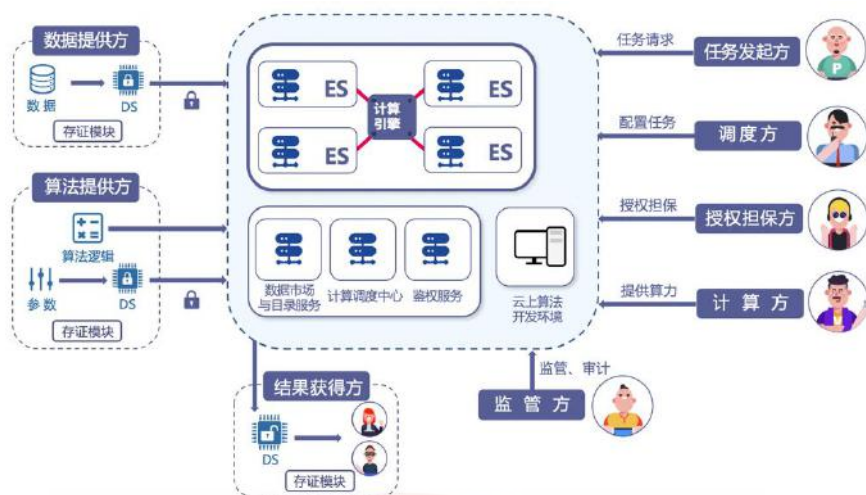


图 16 华控清交多方计算平台应用示意图

平台在数据提供方上部署 DS 模块，将数据转换成计算因子（密文）。平台核心部分是计算引擎，由 4 个计算节点组成，执行密文计算。平台将算力和数据分离、分别独立扩展，这种“管理中心化、信任去中心化”的架构保证了技术应用的安全性和可监管性。

平台通过明密文混合计算框架来支持联邦学习。在该框架下，一个任务的计算过程被拆分成若干明文计算和密文计算，可灵活地进行规划和调度。每个数据提供方部署明文计算引擎和一个密文接入模块 DS：明文计算结果通过 DS 加密后传至密文计算引擎，而密文计算结果也可以通过 DS 解密后继续进行后续的明文计算。在一次计算任务中可以进行明文计算和密文计算的多轮交互。这样联邦学习实际上成为了明密文混合计算的一种应用。

16. 冲量在线平台产品—隐私计算和数据流通产品

冲量在线隐私计算和数据流通产品针对国内主流的芯片平台（包括兆芯、海光、鲲鹏、飞腾等）和操作系统平台（统信、麒麟等）进行了兼容适配，联合硬件和系统共同构建出完整的端到端隐私计算解决方案。帮助企业在自主可控、安全可信的环境中，构建起跨机构进行数据协作的平台，实现数据价值的充分挖掘，同时由对企业机密信息和用户隐私数据进行充分保护，做到数据“可用不可见”。

本产品以基于硬件的 TEE 可信执行环境技术为基础，融合多方安全计算、联邦机器学习、云原生等技术，具有安全可控、完全中立性、交易流程可追溯、技术全面、性能卓越、部署灵活这六大特性，在此基础上通过增加数据集、AI 算法、计算任务、组织账户和日志审计这五大管理模块功能对平台进行补充以便实现更多的功能，能够与区块链系统进行无缝整合，提供金融级的安全计算、加密通信数据检索与处理能力，支持在安全计算平台之上进行安全便捷地跨机构数据流通与生态协作。

在 TEE 技术的适配上，冲量在线根据芯片的特性进行了深度集成，支持 ZX-TCT、FT-TrustZone、Hygon-CSV 等 TEE 方案，并针对国密算法和 TPM 进行了研发了专用的数据信息保密性插件，从而确保了整体方案在安全性上完全自主可控，在政务、金融等涉及核心机密数据的场景具备更强的数据保护能力。

冲量在线为用户提供端到端的隐私计算和数据流通解决方

案，当前该产品与兆芯的合作在 2020 年信通院组织的 “星河” 行业大数据案例征集活动中被评为 “隐私计算标杆案例”。产品系统架构及功能分层如图 17 所示。



图 17 系统架构及功能分层