ICS 35.210
CCS L77

团 体 标 准

T/CCSA 550—2024 T/CAAAD 003—2024

互联网广告 隐私计算平台技术要求

Internet advertising——Technical requirements of privacy computing platform

2024－07－03 发布 2024－10－01 实施

中国广告协会 中国通信标准化协会 发 布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国广告协会和中国通信标准化协会共同提出并分别归口。

本文件起草单位：蚂蚁科技集团股份有限公司、杭州阿里妈妈软件服务有限公司、中国信息通信研究院、深圳市腾讯计算机系统有限公司、北京瑞莱智慧科技有限公司、北京数牍科技有限公司、阿里巴巴（中国）有限公司、北京沃东天骏信息技术有限公司、华为技术有限公司、华为终端有限公司、北京巨量引擎网络技术有限公司、OPPO广东移动通信有限公司、郑州信大捷安信息技术股份有限公司、北京快手科技有限公司、维沃移动通信有限公司、阳狮广告有限公司、上海外国语大学、国家广告研究院、北京勾正数据科技有限公司、百胜中国控股有限公司、秒针信息技术有限公司、北京明略昭辉科技有限公司、华扬联众数字技术股份有限公司。

本文件主要起草人：昌文婷、刘绍国、张晓蒙、郑波、彭晋、孙浩志、窦洪健、杨正军、杨阳、刘金泉、袁鹏程、赵原、宗华、黄楚宇、李克鹏、聂春祺、张开亮、金银玉、单进勇、林战刚、何杰、张泽华、刘璐、欧阳书馨、程勇、落红卫、李世奇、张贝贝、赵乃萱、付艳艳、刘为华、张吉、王昕、张玮、顾明毅、杨燕、崔洪丽、刘景灿、邹文佐、靳亚楠、邢校园、刘建辉、李响、刘崴。

互联网广告 隐私计算平台技术要求

1 范围

本文件规定了互联网广告场景隐私计算平台的整体框架、技术流程、安全要求等内容。

本文件适用于指导互联网广告企业、科技企业、用户机构、第三方机构等，对广告场景隐私计算平台的设计、开发、测试、使用等。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35273:2020 信息安全技术 个人信息安全规范

JR/T 0196—2020 多方安全计算金融应用技术规范

ITU-T F.748.13 共享机器学习技术框架（Technical framework for shared machine learning system）

3 术语和定义

下列术语和定义适用于本文件。

3.1

隐私计算 privacy preserving computation

在保证数据参与方不泄露原始数据和隐私数据的前提下，对数据进行分析计算的一系列信息技术，保障数据在流通与融合过程中的“可用不可见”。

注：隐私计算的常用技术方案有多方安全计算（Secure Multi-Party Computation）、联邦学习（Federated Learning）、可信执行环境（Trusted Execution Environment）等；常用的底层技术有混淆电路（Garbled Circuit）、不经意传输（Oblivious Transfer）、秘密分享（Secret Sharing）、同态加密（Homomorphic Encryption）、差分隐私（Differential Privacy）等。

3.2

多方安全计算 secure multiparty computation

基于多方数据协同完成计算目标的密码技术和协议，实现除计算结果之外不泄漏各方隐私数据。

[来源：JR/T 0196-2020 3.1]

3.3

联邦学习 federated learning

一种使多个参与方在不泄露其原始数据的前提下，能相互协作，来构建和使用机器学习模型的系统或框架。

3.4

可信执行环境 trusted execution environment

基于一定安全需求设计的硬件和软件的组合运行环境，为数据保护提供安全传输、存储和处理等基础服务，保证其上所运行软件的保密性、完整性。可信执行环境通常与富执行环境并存于同一设备，同时运行。

3.5

差分隐私 differential privacy

一种隐私保护模型，用于确保无论输入数据集中是否表示任何特定的数据主体，统计分析输出的概率分布最多相差一个指定值。即保留统计学特征的前提下，去除个体特征以保护用户隐私。

3.6

oCPX 机制 oCPX mechanism

一种以转化成本为优化目的，根据单个流量的点击率和转化率进行智能动态出价的调整，帮助广告主有效的控制转化成本，提升广告效率，最终达成目标的工具。

3.7

需求方平台　Demand Side Platform

需求方平台支持多维度定向手段，用于筛选出最有传播价值的人群，锁定广告目标受众。主要定向方式有：消费意向定向、地域定向等。

注：需求方平台（Demand Side Platform），是为广告主提供跨媒介，跨平台，跨终端的广告投放业务技术产品服务平台，通过确权数据的整合，分析实现基于用户受众的精准化投放，并且实时监控不断优化。

4　缩略语

下列缩略语适用于本文件：

ADX：互联网广告交易平台（Advertising Deal Exchange）

DSP：需求方平台（Demand Side Platform）

IV：信息价值（Information Value）

PSI：隐私集合求交（Private Set Intersection）

ROI：投入产出比（Return On Investment）

RTB：实时竞价（Real Time Bidding）

TEE：可信执行环境（Trusted Execution Environment）

WOE：证据权重（Weight of Evidence）

5　系统架构

5.1　架构图

互联网广告隐私计算平台包含隐私计算引擎层、算法层、广告应用层、跨层的平台管理和安全保障，如图1所示，其中：

——隐私计算引擎层包括多方安全计算、联邦学习和基于可信执行环境的隐私计算；

——算法层涵盖隐私计算实现的算法与功能，包括样本生产与存储能力，样本集合求交能力、匿踪查询能力、联合建模能力、联合预测能力、联合统计能力以及特征工程能力；

——广告应用功能层包括广告投放、受众分析、反欺诈、归因分析等功能；

——平台管理包括用户管理、节点管理、项目协作管理、数据管理和任务管理；

——安全保障包括计算安全、数据安全、通信安全、系统安全和应用安全。

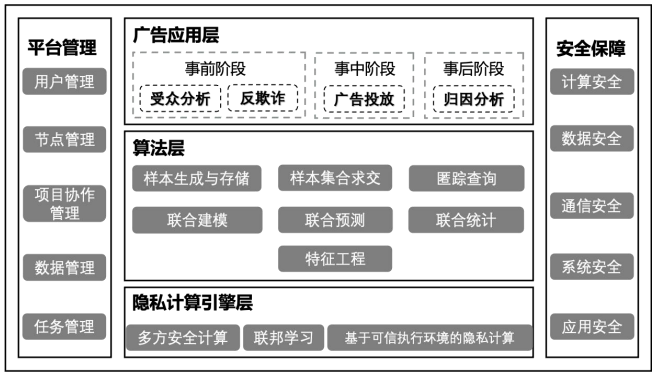

图 1　互联网广告隐私计算平台的总体架构图

5.2　参与方

互联网广告场景下的隐私计算参与方主要包含流量方、广告主和数据方，功能要求和安全要求适用于所有参与方。其中：

a)　流量方：通常是媒体平台，拥有用户前链路行为以及媒体兴趣偏好；

b） 广告主：通常为电商、教育、游戏、旅游等行业平台，广告主拥有用户深度转化链路相关的数据信息；
c） 数据方：除广告主、流量方外的，提供用户数据的参与方，提供多维度的用户偏好信息，辅助广告投放分析，例如数据交易所等。

6 广告应用层要求

6.1 总体说明

基于隐私计算的广告业务流程，可分为事前阶段、事中和事后三个阶段：
a） 事前阶段用于制定广告投放策略，为广告投放过程提供参考依据,包括受众分析，反流量欺诈等业务流程；
b） 事中阶段可根据效果调整投放策略，包括广告投放业务流程；
c） 事后阶段用于广告投放效果归因分析。

根据实际业务场景，隐私计算平台可根据实际广告业务需求，应至少满足6.2、6.3、6.4和6.5所述一种广告应用功能要求。

6.2 广告投放

互联网广告应用在广告投放方面的应用场景如附录A所示，其功能要求如下：
a） 隐私计算平台宜通过整合多方数据构建立体的画像，实现数据资源的优势互补；
b） 流量方宜支持基于自身拥有的大量行为信息和基础画像数据，与广告数据方拥有的深度转化链路数据（如付费信息）进行安全求交；
c） 各参与方应支持通过隐私计算技术联合训练、建模、优化广告投放模型，通过该方式帮助多方在不输出原始数据的基础上共享各自的用户数据进行广告投放模型建模，并根据建模结果制定投放策略；
d） 隐私计算平台应支持对模型效果的监控功能，用于分析并提升模型应用的效果；
e） 隐私计算平台应支持对广告投放效果的分析，用于提升和改进算法，提升广告效果。

6.3 受众分析

互联网广告应用在受众分析方面的应用场景如附录B所示，其功能要求如下：
a） 隐私计算平台宜通过整合各数据参与方，构建受众标签特征体系和受众全景分析画像；
b） 隐私计算平台应支持各参与方在保证原始数据不出本地的基础上，实现多方数据隐私求交和联合统计，丰富特征标签和数据统计维度，为制定广告投放策略提供数据支撑；
c） 隐私计算平台应支持实时多方安全计算，及时更新、优化特征标签和统计维度，指导广告策略调整；
d） 隐私计算平台宜支持对受众分析准确度的评估，用于改进受众分析模型；
e） 隐私计算平台宜支持将受众分析结果输出，以帮助广告主提供精细化的受众人群，辅助广告投放策略。

6.4 反欺诈

互联网广告应用在反流量欺诈方面的应用场景如附录C所示，其功能要求如下：
a） 隐私计算平台宜通过整合多方数据，实现反欺诈数据资源共享互补；
b） 隐私计算平台宜支持多个数据方基于自身大量的反欺诈数据，与广告主所有拥有的；
c） 数据进行深度的安全求交，丰富反欺诈数据维度；
d） 各参与方宜支持通过隐私计算技术联合训练、建立、优化流量反欺诈模型，通过此方式在保证数据安全的前提下，共享广告主、多个流量平台之间反欺诈数据，进行联合建模，根据共同训练的反欺诈模型，过滤注水流量，制定投放策略；
e） 隐私计算平台宜支持对流量反欺诈模型效果的监控、分析，用于提升反欺诈模型效果。

6.5 效果归因分析

广告主在向不同流量方同时投放广告的场景中，对于非直接转化阶段，存在无法确认各个流量方渠道广告效用的问题。基于隐私计算的广告归因分析可结合不同流量方的数据，帮助广告主评估不同流量方广告对转化的贡献份额，辅助制定广告投放策略。

互联网广告应用在效果归因分析的应用场景如附录D所示，其功能要求如下：

a) 隐私计算平台宜结合不同参与方数据，解决单一流量方、广告主方数据不足的问题；
b) 各参与方应支持安全求交功能，构建转化链路信息；
c) 各参与方宜支持通过隐私计算平台进行联合建模和归因分析；
d) 隐私计算平台宜支持对归因分析准确度的评估，用于改进归因模型；
e) 隐私计算平台宜支持将归因分析结果输出，以帮助广告主评估不同流量方广告对转化的贡献份额，辅助制定广告投放策略。

7 算法层要求

7.1 总体说明

根据实际业务场景，隐私计算平台可根据实际广告业务需求，选择性实现下述算法，但应至少满足下述一种算法能力要求。

7.2 样本生成与存储能力

隐私计算的数据提供方，暨持有参与计算数据的参与方，应当保证使用原始数据生成样本过程与样本存储过程的安全性：

a) 样本生成：
   1) 样本生成构建流程应在数据提供方可控域进执行，即在样本生产过程中不能将原始数据、中间数据、特征计算规则、样本分布等信息直接暴露给其它参与方；
   2) 如在特征生成阶段需要进行诸如交叉、匹配、组合等融合方式的，应施加相应的安全保护手段，施加的安全手段应可以阻止对方推断出己方使用的特征域、特征值、特征分布等等。
b) 样本存储：
   1) 数据提供方所构造样本应存储在可控域的存储空间内，存储介质由己方决定，如文本、数据库等；存储方式由己方决定，如单机、分布式等；
   2) 应保障存储内容的安全性，其它参与方在未授权前提下不应有访问存储服务、获取存储内容权限。

7.3 样本集合求交能力

样本集合求交是联合训练的前置步骤，联合训练参与方需要在本阶段获取各自持有样本的交集用于后续训练流程，并且需要生成交集部分数据的描述文件以便后续训练过程中进行样本对齐。样本集合求交流程应保证正确性、高效性、稳定性、求交结果的一致性与样本的安全性。

a) 正确性：样本集合求交流程，需要保证交集结果的正确性；
b) 高效性：样本集合求交流程应采取相关技术手段提升求交过程的效率，包括但不限于采用提高并发、通信压缩、增量求交、改善加密手段等提高计算效率的方式；
c) 稳定性：样本集合求交过程中各参与方应协同保障对齐过程的稳定性，在资源保障、失败任务现场恢复等等方面应给予充足的保障手段；
d) 一致性：
   1) 求交流程应确保求交结果的一致性，即双方特征在同一条样本中关联正确；
   2) 求交流程应确保求交结果对双方键值关联的正确性；
   3) 若有相关业务需求(如双方 key 有重复)，应确保求交结果的唯一性。
e) 安全性：样本集合求交应当在适当的安全保障之下进行，以免将己方隐私泄露给其它参与方。

7.4 特征工程能力

互联网广告隐私计算平台应具备特征工程能力，包括但不限于特征筛选、特征过滤、特征转换、缺失值填充、异常值截断、数据缩放、数据编码、IV值与WOE编码等。特征工程过程应满足以下要求：

a) 安全性：在计算过程中应施加适当隐私计算技术手段以保证计算过程的数据安全，避免各参与方输入数据及中间数据等遭到泄露；
b) 准确性：应保证计算结果精度满足各方预先约定的精度值要求；
c) 高效性：特征工程阶段耗时应满足应用场景需求。

7.5 联合训练能力

联合训练指各参与方使用己方的数据、模型参数在交互流程中共同训练一个全局模型。联合训练作为一种跨域的分布式训练模式，联合训练宜支持深度神经网络模型的训练。此外，联合训练在隐私保护的约束下应当满足以下要求：

a) 安全性：在联合训练过程中涉及到模型中间结果、梯度的传输，而在不加以保护的情况下可能会造成参与者的隐私泄露，因此应在联合训练过程中施加适当隐私计算技术手段以保障训练过程的数据安全；
b) 高效性：联合训练应以适当的技术手段提升模型的训练效率，以保障模型迭代诉求。训练过程宜能够支持千万级的样本量；
c) 稳定性：联合训练过程稳定性应符合以下要求；
   1) 训练参与方应在己方稳定性保障如机器资源、故障恢复、资源等方面做好保障；
   2) 各参与方间在训练故障恢复等方面应做好协调和适配。
d) 灵活性；
   1) 联合训练宜提供更灵活的交互，如单方变更模型需重新进行训练，在对方授权的情况下由己方直接拉起训练任务，减少迭代沟通成本；
   2) 联合训练的参与方各自宜持有训练框架和通信协议，以便在训练流程中可以让新参与方通过简单的协议适配后快速参与到训练中来。
e) 完备性：除训练基础功能外，联合训练流程应当对相应提供模型指标安全透出功能、离线模型转换功能以便在诸如在线交互、跨框架模型继承等场景使用。

7.6 联合预测能力

通过联合训练流程得到全局模型后，各参与方应具备使用联邦模型进行预测的能力。根据是否需要进行在线服务，预测过程可分为离线预测和在线预测：

a) 离线预测：如使用测试集对样本模型进行测试得到模型在测试集上的效果指标，这种情况下可复用数据求交、联合训练链路，获取最终指标；
b) 在线预测：在线预测涉及到各参与方搭建己方的在线预测服务，将己方的子模型交付到在线服务待实时进行预测。

预测过程需满足以下要求：

a) 在隐私保护的前提下，应支持相应的安全保障手段保证过程中数据安全；
b) 参与方应提供模型转化的基本功能，支持离线模型、在线模型的转化工作。

7.7 匿踪查询能力

匿踪查询能力是指数据查询方在请求数据提供方的数据时，无需向数据提供方暴露自己请求的对象的能力。匿踪查询需要被构建成为在线服务，查询方通过调用在线服务的API接口获取查询结果。匿踪查询过程应保证安全性、高效性、稳定性和查询结果的准确性。

a) 安全性：匿踪查询应在密码协议的保护下进行，避免造成查询方暴露请求对象，或数据提供方泄露除查询结果外的数据；
b) 高效性：匿踪查询服务应该兼顾网络传输能力和计算资源，通过技术手段提升查询效率；
c) 稳定性：匿踪查询服务应稳定可调用，服务提供商应在如机器资源、故障恢复、资源等方面做好保障；
d) 查询结果的准确性：应保证匿踪查询的结果与明文查询结果一致。

7.8 联合统计能力

互联网广告隐私计算平台应具备联合统计能力，包括支持方差、最值、分位数等统计运算。联合统计过程应满足以下要求：

a) 安全性：在联合统计过程中应施加适当隐私计算技术手段以保证过程的数据安全，避免各参与方输入数据及中间数据等遭到泄露；
b) 准确性：应保证计算结果精度满足各方预先约定的精度值要求；
c) 高效性：联合统计阶段耗时应满足应用场景需求。

8 计算引擎层要求

隐私计算引擎层应支持多方安全计算、联邦学习、基于TEE的隐私计算等基础算法的实现，如附录E所示，且至少具备下述三种能力中的一种：

a) 应支持多方安全计算，如秘密分享、不经意传输、同态加密等能力；
b) 应支持联邦学习，如分布式本地计算和中心汇聚、广播等能力；
c) 应支持基于 TEE 的隐私计算，如支持在可信执行环境中数据解密和执行计算等能力。

9 平台管理层要求

9.1 总体说明

隐私计算过程中有多方参与，流程相对复杂，因此对隐私和权限管控要求较高，需要参与方频繁协调。因此为了降低用户使用成本，对隐私计算过程的抽象，宜支持使用者可视化地进行用户、节点、项目、数据、任务的管控，提升样本模型实验的迭代效率，缩短开发周期。平台管理能力可以根据参与方角色的实际适用情况，将管理能力设置为必选和可选。

9.2 用户管理

互联网广告隐私计算平台在用户（即各计算参与方的实际操作人员）管理方面功能要求如下：

a) 应支持各参与方的注册及其信息管理功能；
b) 应支持各参与方之间建立并维护配对联合关系的功能；
c) 应支持各参与方定义本方操作人员的角色及其操作权限的功能；
d) 宜支持各参与方对操作人员使用行为进行审计的功能。

9.3 节点管理

互联网广告隐私计算平台在节点（即各计算参与方用于计算的计算机资源抽象）管理方面功能要求如下：

a) 应支持各参与方添加、管理以及删除本方的计算节点；
b) 宜支持各参与方监测计算节点的容量、负载以及故障状态等情况；
c) 宜支持对参与方角色和数量的扩展，确保节点部署灵活和业务可扩展。

9.4 项目协作管理

互联网广告隐私计算平台在项目协同管理方面功能要求如下：

a) 应支持项目参与人员的管理；
b) 应支持项目所使用数据集的管理；
c) 应支持项目参与人员的不同操作权限，应设立项目管理者角色。

9.5 数据管理

互联网广告隐私计算平台在数据（即各计算参与方参加计算后得出的计算结果）管理方面功能要求如下：

a) 应支持数据元信息如存储系统及其地址、操作人及操作时间等的管理和展示；
b) 应支持数据使用和操作过程的授权和权限控制；
c) 宜支持数据使用和操作过程的审计和记录；
d) 宜支持过期数据的清理和误删除数据的恢复；

e） 宜支持模型数据的治理功能，包括多模型的导入、导出、上线等。

9.6 任务管理

互联网广告隐私计算平台在任务（即各计算参与方执行计算过程的最小单元）管理方面功能要求如下：

a） 应支持计算任务的配置、创建、停止等基本操作；
b） 应支持计算任务的运行状态、错误信息等关键信息的查看，对在线任务的调用量、调用成功/失败情况进行监控；；
c） 应支持提交和处理在线模型训练或推理的任务流；
d） 宜支持提交和处理离线批量模型训练的任务流；
e） 宜支持计算任务与其他参与方相关联任务（如有）的关联关系维护；
f） 宜支持任务排队、多任务并行、任务历史信息等增强扩展功能。

10 安全要求

10.1 计算安全

互联网广告隐私计算平台在计算安全方面的要求如下：

a） 应保证各隐私计算技术的安全模型与安全参数满足应用场景的需求；
b） 应采用同态加密、秘密分享、差分隐私等技术手段，保证中间数据传递、融合过程中的安全，如使用联邦学习方案时梯度融合的情况；
c） 宜采用相应技术手段，保障计算环境安全。

10.2 数据安全

互联网广告隐私计算平台在数据安全方面的要求如下：

a） 计算过程中，参与方应只能获得其参与计算所必需的数据和协议规定的计算结果，不能获得其他任何信息；
b） 应采用技术措施保证隐私计算过程中不出现其他参与方的隐私数据；
c） 应加强各节点的隐私保护能力，不因单点故障而泄露任何一方相关信息；
d） 应保证隐私计算的结果数据安全；
e） 应采取有效措施对数据使用过程进行管控，并采用安全、不可篡改的方式对中间、结果数据进行存证。

10.3 通信安全

互联网广告隐私计算平台在通信安全方面的要求如下：

a） 应采用安全的协议保证隐私计算各参与方的通信安全；
b） 在通信节点建立连接之前，应进行通信方双向身份认证；
c） 应采用密码技术对通信数据进行保密性和完整性保护；
d） 当通信数据被篡改后，数据接收方应能识别并立即采取措施进行异常处理。

10.4 系统安全

互联网广告隐私计算平台在系统安全方面的要求如下：

a） 应采取一定的技术手段保障执行环境系统安全；
b） 宜采取白名单控制策略，并使用隐藏或掩码方式防止侧信道攻击；
c） 系统宜具有防 DDOS 攻击等安全防护能力；
d） 应对登录隐私计算平台的用户分配账户和权限，进行身份标识和鉴别；
e） 参与方的用户身份标识应具有唯一性，身份鉴别信息应具有复杂度要求并定期更换。

10.5 应用安全

互联网广告隐私计算平台在应用安全方面的要求如下：

a） 应对应用过程中的关键操作进行身份认证；

b) 应采取技术措施保证操作行为的合法性和抗抵赖性，避免出现操作风险；
c) 多参与方进行计算任务时，各参与方宜支持对计算任务进行审批。

附　录　A
（资料性）
广告投放场景

A.1　基于联邦学习的广告投放

目前联邦学习技术已经在金融领域大规模应用，在广告搜索推荐这种大规模稀疏场景领域的应用和研究尚处于发展初期。伴随着外部媒体短视频流量异军突起，商家有从媒体引流电商的需求，而商家在媒体直投存在后链路效果分析成本高、无法同时在多个媒体投放的问题。而且伴随着越来越强的隐私监管，企业数据互为商业机密，无法直接共享。基于联邦学习的广告平台可以在隐私保护的前提下，通过个性化建模提升商家投放广告的效果优化诉求。

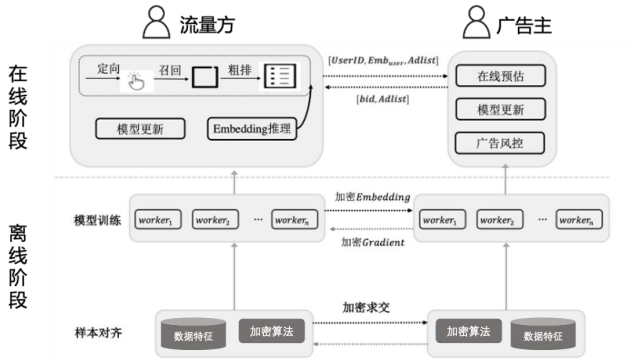

图 A.1 基于联邦学习的广告投放场景框架图

参与方主要包含流量方和广告方。流量方通常是媒体，拥有用户前链路行为以及媒体兴趣偏好等，而广告主通常为电商、教育、游戏、旅游等行业平台，广告主拥有用户深度转化链路相关的数据信息。业务流程分为在线阶段和离线阶段。

在线阶段的业务流程如下：

a) 流量方、广告主、第三方数据提供方会在严格保护其各自数据隐私的前提下，基于联合训练的点击转化率预估模型和 oCPX 机制为用户推荐感兴趣的广告，以保证用户体验和广告主的广告投放效果；

b) 用户根据兴趣点击后，将跳转到广告主的业务平台，广告主的业务平台会存有商品特征、用户历史特征以及本次点击的信息等。

离线阶段的业务流程如下：

c) 流量方、广告主、第三方将在线模型产生的相关信息生成样本数据；

d) 随后三方将采用基于加密的样本集合求交将加密的样本数据对齐；

e) 样本数据对齐后，流量方与广告主可以选择直接将对齐的样本数据结果进行分析，确定投放策略或将采用对齐的样本数据，进行模型训练，通过模型训练结果确定投放策略。在训练过程中，有标签的广告主作为主导方，协同方为流量方、第三方。

1) 协同方将一个训练的中间结果经过加密之后发送给主导方；

2) 主导方在计算中间数据后，将协同方发送来的中间结果对应的梯度经过加密后发送给协同方，完成训练迭代；

3) 最终实现在保护数据隐私的情况下，流量方与平台方共同实现模型训练。

经过联邦建模分析及模型优化，显著提升商家ROI的效果，且帮助商家实现数据闭环的全链路优化。

A.2　基于多方联邦学习的广告投放

广告是互联网企业的主要变现手段之一，它的核心是利用数据挖掘用户历史行为对用户的兴趣进行建模，然后提供个性化的广告，从而提高用户的广告体验，降低广告主的投放成本，并且提升平台的收益，达到多方共赢的局面。

大数据是人工智能时代的石油，但是由于监管法规和商业机密等因素限制，"数据孤岛"现象越来越明显。在广告场景中，流量主、数据方和广告主侧各拥有一部分链路数据，不能完全同步给对方，但是双方都有需求优化广告投放效果，以提升成本控制和起量效果。借助多方联邦学习可以在保护合作各方各自数据安全的前提下，联合训练、建模、优化模型效果。

在这样的背景下，通过广告主、数据方和流量主的多方联邦学习计算，融合各方的数据优势，在广告应用案例中能够取得显著效果提升。

基于多方联邦学习的广告投放的框架图如图A.2所示：

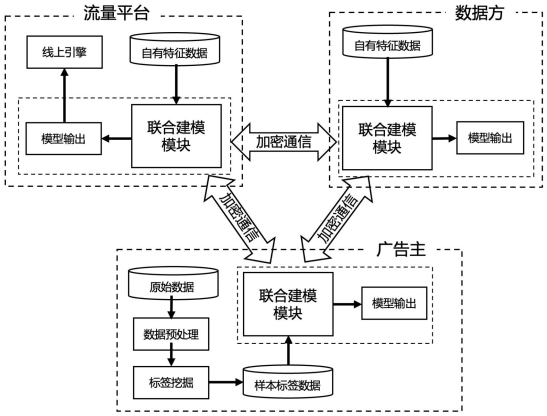

图 A.2 基于多方联邦学习的广告投放的框架图

在本应用场景中，包括广告主、流量平台和数据方三个参与方。这三个参与方中，均部署有联合建模模块。数据方和流量平台上，各自拥有自有特征数据，基于联合建模模块，实现模型输出。广告主一侧，通过对原始数据的预处理、标签挖掘，提炼样本标签数据，通过与流量平台和数据方的联合建模，各自得到模型输出。

在本应用场景中，广告主通过过往广告投放转化的高质量样本，基于数据方和流量平台的数千纬度的特征，进行多方联合建模。在完成模型训练后借助平台的隐匿查询功能，可以完成批量设备号联邦预测打分。

在整个过程中，各合作方的原始数据始终不出本地，平台通过交换加密中间参数来完成联合建模，提升广告投放的效果。

A.3　基于多方联合建模的 RTB 广告投放

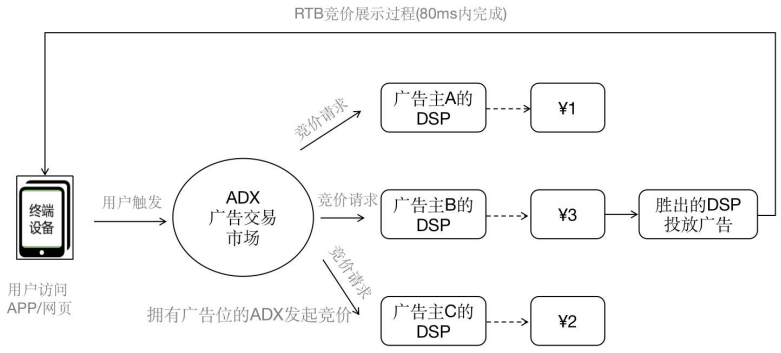

图 A.3 RTB 实时竞价流程图

在RTB广告实时竞价中，广告主的DSP虽然结合了海量的用户信息进行了用户分析以及定向投放，然而，广告主手中也同样具有影响广告投放的重要用户信息，例如：用户的活跃度、购买服务等。为了满足广告主实时个性化的投放需求，RTA将广告的流量选择权交给广告主。RTA在定向环节中将用户身份的识别的请求发送给广告主，进行用户的筛选，让广告主在广告曝光前进行投放策略的判断，满足“拉新”、“拉活”等个性化需求，如图A.4所示。具体来说，在RTA对接服务中，DSP平台需要将预投放的用户发

送到广告主端，广告主端再通过隐私求交（PSI）技术过滤掉不符合自己定向的客户。同时，也可以在RTA中引入安全联邦学习技术利用双方数据进一步提高双方数据的利用价值，使广告主高效获客，从而实现最大化广告主和平台的利益。

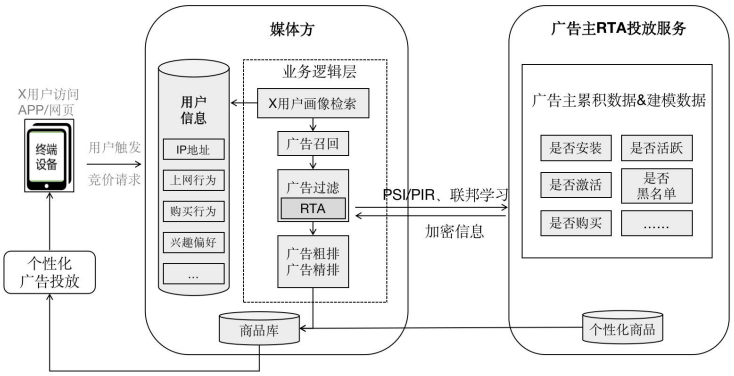

图 A.4 基于 RTA 的广告投放流程图

RTA本质来说，是解决广告系统平台无法实时个性化定向的一个需求。

a) 广告主具备一方数据，但是投放的目标人群实时变动，通过平台的定向标签能力无法实现精准定向；又或者通过用户包无法实现实时定向更新，需要结合双方的数据能力共同提升广告主投放效果。RTA 激活了广告主的用户甄别和筛选能力；

b) 出于数据安全或者价值的考虑，广告主不愿意将转化数据回传给广告平台。比如金融公司投放金融广告的时候，需要将无效征信的用户去除，但是由于无效征信的属于高度敏感的数据，金融广告公司出于数据安全的考虑，无法将数据传到广告平台上。所以 RTA 实现了保障广告主的隐私数据不出库的情况下的用户筛选；

c) 广告主需要个性化买量需求，针对不同的用户选择不同的投放策略和方案。比如，不同的公司的增长团队，对于纯新增用户，安装卸载，安装不活跃用户（一周内未打开过 APP 且使用了某个功能）有不同的投放策略。很多精细化的数据逻辑只有广告主有，这个时候可以通过 RTA 进行广告的个性化投放。

附　录　B
（资料性）
受众分析场景

B.1　基于多方联合统计的受众分析

广告主在广告投放之前，需要结合企业的营销诉求，经过海量收集、关联整理、链路分析目标受众的信息，从而精准定位受众需求，最后通过个性化推送广告活动实现营销目的。受众分析是广告投放的前提，直接影响广告投放效果。

但广告主能收集和整理的受众信息，仅限于由流量主提供的、因广告合作而产生的广告群体行为数据，以及广告主针对个体用户运营过程中产生的长期兴趣行为数据。广告主亟需打通第三方数据源，结合样本数据拓展更多的目标受众，补全更多维度的受众标签，俯瞰全行业的受众群体画像。

在此背景下，基于多方联合统计，可以打通广告主和第三方数据源各自拥有的受众信息，在保护合作各方数据安全的前提下，来联合统计受众信息，从而构建完备的受众标签特征体系和受众全景画像，完成精准的受众需求洞察，提升广告投放效果。

基于多方联合统计的受众分析框架图如图B.1所示：

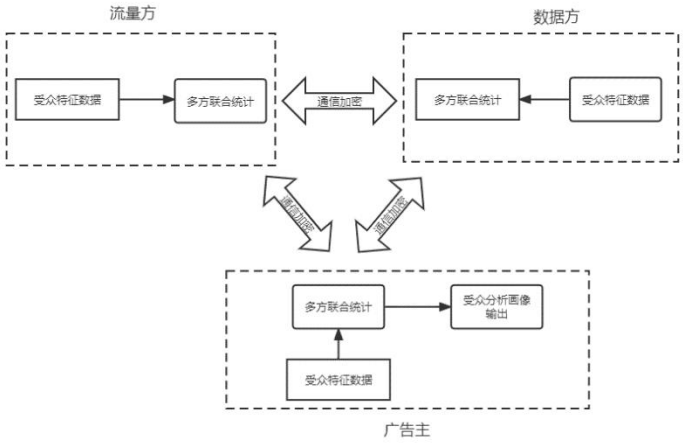

图 B.1 基于多方联合统计受众分析的框架图

在本应用场景中，包括广告主、流量主和数据方三个参与方。这三个参与方中，流量主可以通过广告投放后台的数据报告，将受众群体信息供给广告主。给广告主和数据方部署隐私计算平台，然后基于匿踪查询组件，实现两方数据的联合建模，模型输出给广告主的同时保证广告主的受众信息不被数据方可知。

在本应用场景中，广告主通过多方安全计算技术，安全地构建标签特征体系和全景画像，为精准的受众分析提供数据支撑。从而保证了广告投放事前，其账户搭建、受众定向、素材撰写、广告活动设计、转化路径设计等广告策划更加精准化、个性化，即提高了受众用户的广告体验，又提升了广告主的投放效果。

B.2　基于联邦建模的受众分析

场景描述：广告主在广告投放平台上投放广告，不同的转化目标，会有不同的转化人群，在投放过程中，对应不同的受众人群。

a)　对于广告主的深度目标，例如付费、7 日用户留存等深层转化目标，对于这类任务，广告主的主要目标是唤醒用户，刺激用户的付费等行为，广告主会有站内的用户行为，例如用户浏览内容、加购商品行为、用户付费商品内容等等用户在站内的行为数据，这类数据通常是长周期的，能够直接反映用户的意图，兴趣以及习惯，广告投放平台是没有这部分用户特征的；

b) 在互联网广告中，广告主通过 DSP 来完成广告投放任务，DSP 在广告主投放任务时，为广告主提供目标受众人群数据，通常包括：地域偏好、年龄偏好、性别偏好、学历偏好、消费偏好等。

当前，广告投放平台提供的群体数据维度单薄，粒度较粗，范围较大，对于不同的广告主，通用的受众定向不能很好的满足特定广告的受众用户，换句话说，对于不同的广告需要精准的找到符合某一广告主的用户，才能更好地提高广告主的深度转化目标。

在此背景下，基于多方联邦建模，可以打通广告主和广告投放平台方各自拥有的部分受众信息，在保护合作各方数据安全的前提下，来联合广告投放平台的广告表现数据以及广告主方的用户长期意图兴趣数据，构建广告主的特定广告任务的受众人群、非受众人群，完成精准的受众需求洞察，提升广告投放效果。

基于多方联邦模型的受众分析框架图如图B.2所示：

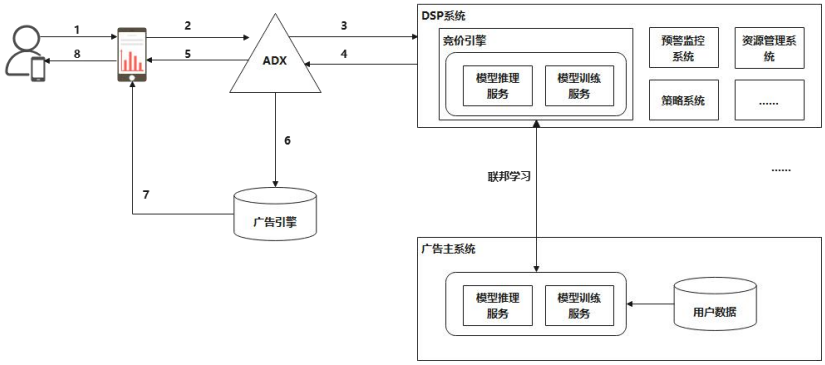

图 B.2 基于多方联邦模型受众分析的框架图

在本应用场景中，包括广告主、广告投放平台两个参与方。这两个参与方中，广告投放平台可以通过DSP系统中的广告投放数据，结合广告主系统中的用户深度转化的表现数据进行联邦学习，为广告主生成特定的受众人群与非受众人群，联邦学习的结果可以指导人群投放策略，更好的进行广告任务的投放，提升转化效果。

a) 用户产生广告请求；
b) 媒体将携带客户标识（一般是 Cookie 或设备号）的流量发送到媒体 ADX；
c) 媒体 ADX 向广告主指定的 DSP 发起曝光竞标请求；
d) DSP 在满足 N 倍推送比例的约束下，进行估值后决定是否选择本次流量，将结果返回给媒体 ADX；
e) 如果 DSP 选择本次流量，ADX 按照媒体广告模板进行样式渲染后，将 DSP 的广告返回给客户展示；
f) 如果 DSP 没有选择本次流量，ADX 将流量重新返回给广告引擎；
g) 广告引擎重新选择其他广告后，返回给客户展示；
h) 客户浏览页面，看到广告，广告产生曝光。

在本应用场景中，广告主和广告投放平台体系两盒建模，可以实现受众分层以及需求洞察，针对不同的广告投放任务可以细化每种广告任务的用户人群，帮助广告主更加精细化、精准化的投放广告。为受众推荐个性化广告的同时，即提高了用户的广告体验，又提升了广告主的投放效果。

附　录　C
（资料性）
反欺诈应用场景-基于多方联合建模的反欺诈应用

当前流量作弊方式更新快、拦截难度大。广告代理商和广告交易平台，引入虚假流量，以次充好；黑产方，提供自动化的工具、专业刷量团队、薅羊毛团队，恶点流量。面对买假量、刷假量、掺假量等流量作弊问题，广告主需要在媒体方RTA（或DMP平台）上部署异常流量判断模型，从而在广告投放事前，对个体流量做有效性的实时判断。

广告主仅能收集流量承接平台上的个体访客基础信息和群体浏览行为信息，信息维度单一，亟需打通第三方反欺诈数据源，构建更科学的反欺诈模型。

在此背景下，借助隐私计算平台，在保护广告主、数据方数据安全的前提下，通过多方联合建模，构建反欺诈模型。在广告投放的过程中，通过RTA平台和DMP平台实时比对流量的有效性，最终排除可疑流量、提升广告效果。

基于多方联合建模的反欺诈模型的框架图如图C.1所示：

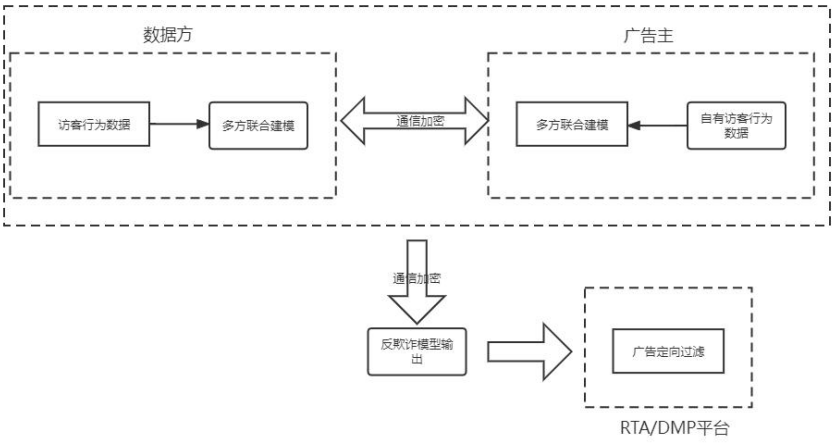

图 C.1 基于多方联合建模的反欺诈模型的 RTA 场景下的框架图

在本应用场景中，包括广告主、数据方、RTA（或DMP）平台、流量主4个参与方。广告主一侧，通过对访客行为信息数据的预处理、标签挖掘，提炼样本标签数据，数据方一侧，拥有自有特征数据。为两个参与方部署有联合建模模块，对广告主端实现反欺诈模型的输出。输出的模型作用于RTA（或DMP）平台，从而完成流量主流量的实时判断，最终决定该流量的竞价策略。

在整个过程中，广告主和数据方的原始数据始终不出本地，平台通过交换加密中间参数来完成联合建模，再通过RTA（DMP）平台做流量的实时判断，从而实现了广告投放的前置异常流量过滤，节省了广告成本。

附　录　D
（资料性）
效果归因分析场景-基于多方联合建模的归因分析

广告主使用全域营销策略，通过多个流量平台，全方位的触达客户，提升品牌声量，抢占市场份额。

广告主做归因目的是将消费者的最终转化，合理分配给转换途径中的各个平台触点，从而妥善衡量各渠道广告的成效，并针对后期投放制定最佳决策。

广告主构建广告归因分析模型，需先打通多个流量平台的消费者互动信息数据。

在这样的背景下，通过广告主、各方流量主的多方联合建模，构建归因模型，从而更好地了解广告的效果，并针对转化链路中的各个广告投放阶段采取对应的优化措施。

基于多方联合建模的归因分析的框架图如图D.1所示：

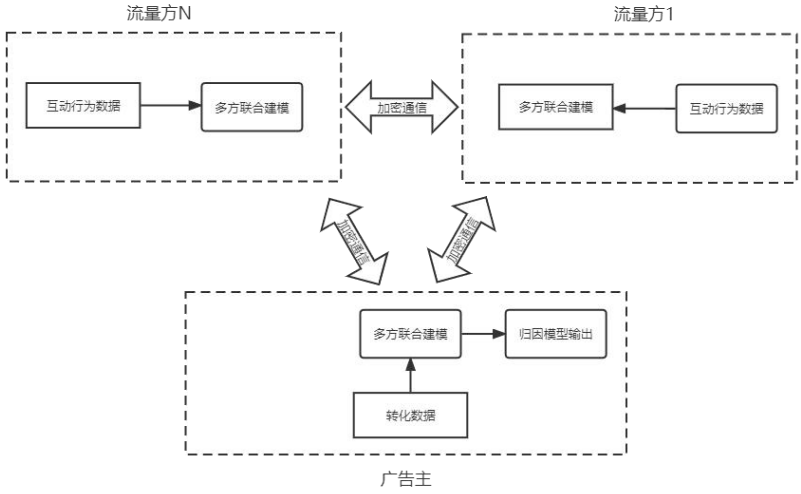

图 D.1 基于多方联合建模的归因分析的框架图

在本应用场景中，包括广告主、N方流量平台等多个参与方。这多个参与方中，均部署有联合建模模块。广告主一侧拥有原始转化数据，流量平台方拥有互动数据，然后通过多方联合建模，最终输出归因模型给广告主。

广告主通过使用归因模型，妥善衡量各渠道广告的互动过程成效和互动转化成效，并优化广告投放预算的分配、出价策略的调整、投放力度的优化，为广告投放制定最佳决策，提升了全域营销整体效果。

附　录　E
（资料性）
隐私计算技术

安全多方计算（MPC）是指在无可信第三方的情况下，多个参与方协同计算一个约定的函数，并且保证每一方仅获取自己的计算结果,无法通过计算过程中的交互数据推测出其他任意一方的输入和输出数据。包括同态加密、秘密共享、混淆电路、零知识证明、不经意传输（OT）等技术。安全多方计算属于分布式加密计算，一般不涉及到机器学习联合建模，多用在匿踪查询、统计分析、多方协同运算等场景，他们主要通过生成并交换随机数据实现隐私保护，并通过预先设计的计算协议保证计算结果的有效性，这种典型思路适合进行精确计算和数据库的查询操作，并且能够证明其计算安全性。

联邦学习又名联邦机器学习、联合学习、联盟学习。联邦机器学习是一个分布式机器学习框架，它将传统的机器学习进行定制化的隐私保护改造，能够在隐私保护的基础上帮助多个机构，在进行数据合作及联合建模时，有效地保护用户隐私及数据安全。其中，联邦学习又分为：横向联邦学习、纵向联邦学习、联邦迁移学习。

基于可信执行环境的机密计算是一种将敏感应用和计算业务运行于可信执行环境中,以达到更高安全保障的计算模式。目标是防止敏感数据以及代码的泄露和滥用，保证敏感数据的保密性、完整性，以及代码的完整性、可信性。为了达到更强的纵深安全防御，机密计算从技术上为数据安全和数据隐私保驾护航，将大大激发数据的价值，成为数据的应用以及融合创新的新引擎。基于可信执行环境的机密计算系统涉及计算硬件及软件，其中可信执行环境的可信是构建隐私计算系统可信的基础支撑。机密计算系统可以从隔离、加密、远程验证等维度来保证敏感数据及代码的可信和安全。

参 考 文 献

[1] JR/T 0196-2020 多方安全计算金融应用技术规范
[2] ISO/IEC 2382, Information technology — Vocabulary, 2015.
[3] ITU-T F.748.13, Technical framework for shared machine learning system, 2021.