

隐私计算通信应用 研究报告

(2022年)

隐私计算联盟 2022年12月



版权声明

本报告版权属于隐私计算联盟、中国信息通信研究院云计算与大数据研究所,并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的,应注明"来源:隐私计算联盟、中国信息通信研究院云计算与大数据研究所"。违反上述声明者,本院将追究其相关法律责任。





编写委员会

❖ 主要编写单位(排名不分先后):

隐私计算联盟、中移动信息技术有限公司、联通数字科技有限公司、 天翼电子商务有限公司

❖ 参与编写单位 (排名不分先后):

深圳市洞见智慧科技有限公司、京东科技信息技术有限公司、杭州趣链科技有限公司、杭州锘崴信息科技有限公司、上海零数众合信息科技有限公司、同盾科技有限公司、浙江吉利数字科技有限公司、中兴通讯股份有限公司

❖ 编写组主要成员 (排名不分先后):

王思源	闫 树	袁 博	杨靖世
贾 轩	白玉真	童锦瑞	宋佳楠
魏凯	姜春宇	吕艾临	毕剑锋
张帆	茹志强	崔玲龙	闫 龙
李大中	章庆	徐潜	余文青
薛婧	杨辉	杨博	孙中伟
徐静	陆一文	唐丹叶	李 帜
兰春嘉	杨珍	黄翠婷	陈涛
李晨龙	吴 凯	张再军	黄峥

前言

我国高度重视数据安全流通技术的发展应用,在过去的一年内多个部门密集出台了一系列战略、规划和政策,强调数据要素流通的重要性,提出数据安全流通的建设方案。

通信行业为人与人之间的信息交流、传递提供了媒介,在信息化高速发展过程中积累了海量的通信数据,这些数据覆盖全面、特征丰富、真实性高、数据连续性高,蕴含着极高的应用价值和应用潜力。然而,通信数据中也包含了大量敏感信息,在现行的《网络安全法》、《数据安全法》、《个人信息保护法》等一系列法律法规要求下,传统数据流通模式难以同时兼顾数据价值释放与数据安全保护。

隐私计算技术作为保障数据安全流通的有效方式,提供了"数据可用不可见""数据可控可计量"的数据服务新范式,在保障数据安全前提下实现了数据流通效果,从而为需求方企业安全地获取和利用外部数据提供了技术可能。以隐私计算技术为依托,推动通信行业数据与各行业共享赋能,可以在风险管控、营销分析、态势洞察等多个场景提供数据共享服务。在保证数据安全的前提下,充分发挥通信数据应用价值,助力各行业数字化发展实践。

隐私计算联盟联合通信运营商及业内相关技术企业共同完成了本报告的编写工作。报告对隐私计算在通信行业的应用特性及典型场景进行全面梳理,深入挖掘潜在的创新应用场景,为行业发展提供参考指引。

目 录

一、	隐私计算发展背景	1
二、	隐私计算通信应用特性	2
	(一) 通信数据特征及价值	2
	(二) 隐私计算通信应用特性	4
三、	隐私计算助力通信数据价值释放	6
	(一) 赋能金融	6
	(二) 赋能政务	16
	(三) 赋能其他行业	22
四、	通信行业隐私计算创新应用	26
	(一) 智慧城市	26
	(二) 云边协同	28
	(三) 算力网络	30
五、	总结与展望	32
参考	宇文献	34

一、隐私计算发展背景

近年来,数字经济已成为推动我国经济增长的主要引擎之一。国内数字经济规模由 2017 年的 27.2 万亿元增至 2021 年的 45.5 万亿元,总量稳居世界第二,年均复合增长率达 13.6%。在数字经济时代下,万物互联,各行各业的一切活动和行为都将数据化。2019 年 10 月,党的十九届四中全会首次从国家发展战略高度,将"数据"定位为新型生产要素。2020 年以来,中央政策文件中多次强调培育数据要素市场,推动数据要素市场化配置。现如今,数据要素市场建设已成为政策布局和产业发展的关注重点,是促进数据自主有序流动、提高配置效率、发挥数据价值的关键环节。

然而,数据中往往含有大量敏感信息,大规模数据的收集、共享和发布等操作存在泄露隐私数据的风险,给用户带来困扰,甚至危及社会利益和国家安全。近年来,全球各地逐渐提出一系列法律条例来管控组织机构对用户隐私数据的使用,这让以往跨机构直接共享数据的计算模式不再可行。聚焦国内动态,仅 2022 年第一季度,国家层面就发布了多项政策来对个人隐私数据采集、传输、使用、监管等多方面进行了规定。2022 年1月12日国务院发布并实施的《"十四五"数字经济发展规划》中明确指出,要进一步强化个人信息保护,规范身份信息、隐私信息、生物特征信息的采集、输出和使用,并且要求加强对收集使用个人信息的安全监管能力。中国人民银行也正式发布了《个人金融信息保护技术规范》,从安全技术和安全管理两个方面,对个人金融信息保护提出了规范性要求。此外,近年来我国密集出台

实施的《中华人民共和国网络安全法》、《中华人民共和国数据安全法》,《中华人民共和国个人信息保护法》等数据安全保护相关法案,都为隐私计算技术的立法和监管支持奠定了坚实的基础。

隐私计算技术作为保障数据安全流通的有效方式, 乘时乘势高速 发展,已逐渐成为促进数据要素跨域流通和应用的核心技术。产业需 求快速增长,隐私计算走出学院派与实验室,广泛与行业应用场景相 结合,赋能数据价值的安全、合规流转。各类隐私计算厂商也如雨后 春笋一般不断涌现,激发了隐私计算技术可用性的快速提升。通过对 原始数据加密、去标识化或假名化处理, 计算过程及结果只传递经处 理后的数据,隐私计算技术实现了原始数据不出域,保证了原始数据 持有权不变且不受损,仅让渡了数据使用权,实现了数据的持有权和 使用权相互分离,保障了数据主体的合法权益。另一方面,隐私计算 通过限定数据用法、用量,解决了原始数据无限复制、盗用、滥用的 问题。在保护数据要素流通中各参与方的合法权益的同时满足了数据 要素安全流通使用的需求,确立了数据产权配置的全新路径,逐渐成 为促进数据要素跨域流通和应用的核心技术,广泛应用于金融、政务、 通信、医疗、能源等诸多领域。

二、 隐私计算通信应用特性

(一) 通信数据特征及价值

我国通信行业在高速发展过程中积累了海量的通信数据,这些数据覆盖全面、特征丰富、真实性高、数据连续性高,蕴含着极高的应用价值和应用潜力。在数据覆盖范围方面,我国目前手机持有率高达

96%,截至 2022 年 7 月底,三家基础电信企业的移动电话用户总数达 16.72 亿户。由此可见,通信运营商的海量数据有着极高的用户覆盖度。在数据采集深度方面,截至 2022 年 8 月底,中国移动互联网用户数达到 14.59 亿户,用户日均使用时长达 7 小时,月均使用 218.1 小时。用户产生的通信数据不仅包括用户基本信息、通话行为,随着手机终端在移动互联网中的高频使用,通信数据进一步覆盖到用户的上网行为、消费行为、生活轨迹等各个维度。在数据真实性方面,我国当前已全面实施了手机号码实名制,通信运营商可采集到的通信数据具有较高的真实性,能够客观地反映出用户各方面的特征。在数据连续性方面,通信运营商采集的通信数据记录周期长、留存时间长,覆盖了用户从入网到离网全生命周期的海量数据。

结合通信数据在各应用场景中落地的具体情况,对通信数据做出了以下的分类和梳理(如图1)。根据数据源的不同可以将通信数据划分为用户数据和业务数据两大类型。用户数据包括了 IP 地址、入网设备信息等终端信息,姓名、年龄、生日等实名身份资料,区域人口实时密度、生活轨迹等位置数据,网站及 APP 浏览记录、浏览次数、浏览时长等上网行为数据,通话和短信记录、常用联系人等社交关系数据,电话、网络缴费等通信消费数据。业务数据包括了宽带、手机等业务套餐数据,客户咨询与客服回复等客服数据,企业入网信息等政企数据。

数据源	数据类型	数据细分	应用价值	备注
业务数据	业务套餐	宽带、手机套餐	联合互联网企业对内进行5G权益套餐精准营销	
	客服数据	客户咨询与客服回复	提供相关企业用于智能机器人、智能客服系统搭建	
	政企数据	企业入网信息	给金融机构贷款提供企业征信信息	
用户数据	终端信息	ip,入网设备信息	联合公安、人行、金融机构进行黑名单共享、通信反诈、	
	身份信息	用户实名身份证、性别等	金融风控;	联合其他运营商进行5G基 站共建共享
	位置数据	行为轨迹	人口流动密度测算,集合企业工商信息进行客流分析、商 铺选址;联合CDC、公安部分用于疫情防控追踪;	
	上网行为数据	浏览的网站或APP,浏览次 数,浏览时长等	衍生常用网站和APP等偏好信息的用户画像,联合电商企业精准营销	
	通话、短信数据	通话次数,通话时长,短信 发送记录	挖掘常用联系人数据等社交关系数据,用于反诈和人员关 系知识图谱构建,进而联合互联网精准推荐	
	通信消费	用户电话、网络缴费用于信 用分析	给金融机构贷款和理财产品提供用户信用分析	

图 1 通信数据分类及用途

通信数据可被广泛应用于金融、政务、互联网、医疗等行业,丰富的数据特征有助于企业构建更加完善的用户画像,增强风控、营销等业务模型的准确性,进而提升实际业务效果。然而,上述通信数据中包含了部分个人敏感信息,如用户身份信息、位置数据、上网行为和政企数据等,根据现行的《网络安全法》、《个人信息保护法》、《数据安全法》等法律法规的相关要求,这些信息不能直接明文参与计算。因此,在传统数据流通共享模式下,难以同时兼顾数据价值释放与数据安全保护。随着隐私计算技术的成熟,针对这一问题也开始逐渐有了新的解决思路。

(二) 隐私计算通信应用特性

数据作为通信行业的关键性生产要素以及战略资源,推动安全的通信数据市场流通使用,对推动社会经济发展和满足国家对于数据要

素安全合规的要求具有重大意义。以隐私计算技术手段为依托,推动通信行业数据与各行业共享赋能,可以在风险管控、营销分析、态势洞察等多个领域提供数据共享服务,发挥通信数据应用价值,助力各行业数字化发展实践。隐私计算通信应用具有以下特点:

应用数据丰富。通信数据中的营销类相关指标,如网购偏好、消费意愿、理财产品购买偏好、消费等级及活跃程度等信息,能够帮助其他企业构建完善的客户画像,同时提高营销意愿评分判断准确率。通信数据中的风控类相关指标,如用户的通话行为数据、入网时长数据、信用分相关数据等,可以为金融信贷业务等提供良好风控指标数据,企业维度的信用违约信息、号码活跃程度、归属地差异等,能够帮助银行判断小微企业信贷风险。

应用行业广泛。通信数据常被采用对外赋能的形式,解决其他行业数据孤岛问题。在金融领域中,银行、证券、保险等金融机构都有广泛的通信数据联合应用案例。在政务领域中,借助通信数据维度广、数据跨度大等优势,结合隐私计算技术,能够在保障政务数据安全的基础之上,有效发挥多方数据价值。此外,通信数据也被广泛应用于互联网、医疗、汽车等行业的实际业务场景中。

应用场景集中。通信数据被各个行业广泛运用,具体应用场景则相对集中,主要包括联合风控、联合营销和态势洞察分析等场景。联合风控场景的通信数据应用,集中于信贷领域个体用户或企业相关的联合风控、反洗钱与反欺诈场景等;联合营销场景的通信数据应用,集中于潜在客户唤醒营销、金融信用卡客户评分及触达、营销画像构

建等场景;态势洞察分析场景的通信数据应用,集中于政务领域的群体聚集性统计、区域态势信息分析等场景。

应用效果显著。基于隐私计算技术,通信数据在多个行业都有丰富的应用案例体现。在联合风控场景中,充分应用通信运营商的底层数据字段,有效发挥多方数据融合价值,提升了风控模型的精准度。在联合营销场景中,借助通信运营商的大数据优势,帮助企业共同构建更加全面的客户画像,快速准确地筛选目标客群,减少无效的营销投放,提高营销精准度,节约大量成本。

三、隐私计算助力通信数据价值释放

通信数据的应用集中于赋能外部行业领域。通信领域主要作为数据提供方参与隐私计算应用,基于隐私计算技术解决其他行业数据孤岛问题,以"数据可用不可见"的方式为金融行业、互联网企业、政府公共机构等提供合规应用通信数据解决实际问题的途径。

在应用层面,业内开展了大量试点实践,探索应用场景落地可行性并持续推进通信数据智能生态合作落地。通过联合各行业积极探索隐私计算应用场景,不断升级技术安全验证方法及测试标准。在跨行业融合层面,以"数据可用不可见"的方式为金融、政务、电商等行业客户实现与通信数据的融合协同应用提供数据安全保障。

(一) 赋能金融

1.银行信贷智能风控

银行信贷服务业务的核心是进行风险控制。风控管理覆盖信贷业

务的贷前、贷中、贷后全流程,内容涵盖了从贷前准入、信贷申请反 欺诈、信贷额度审批、贷中风险监控、贷后风险预警等。通过构建风 控模型进行信贷申请的风险评估是银行信贷服务业务风险控制的有 效手段,风控模型的质量和成效将会对金融信贷业务的收益产生重要 影响。随着数据量级颗粒度、深度的增加,银行对风控数据模型的需 求、应用都在不断的扩大。然而,当前监管机构对数据使用合规等方 面的约束不断加强,风控模型可用数据源的供给渠道在不断地收缩; 同时,在信贷服务场景中,信息欺诈和数据失真情况不断加重,传统 的银行风控模型因缺乏多维动态数据支撑,已经无法有效识别风险及 进行风险预警。

通过在通信运营商与银行之间搭建基于隐私计算的数据安全共享及联合建模平台,使银行能够安全地利用运营商的动、静态数据,如通话行为标签、入网时长、信贷分、信用卡分等,为业务提供辅助决策支撑。基于隐私计算技术实现的多方联合建模,在各方数据都不出私域的前提下,充分应用通信运营商更多的底层数据字段,并有效发挥多方数据融合价值,提升银行风控模型的精准度,如图 2 所示。



图 2 基于隐私计算的智能风控体系

个人信贷风控方面,借助隐私计算技术,在保障数据安全的前提下融合业务欠费情况、理财偏好、工作稳定程度、阅读偏好等通信运营商数据以及银行自有数据,解决个贷业务场景中信用黑户、多头借贷、贷中逾期、坏账、呆账等风险环节的监测及预警,为银行、信贷、电信、保险等业务活动提供强有力的个人客户风险防控保障。

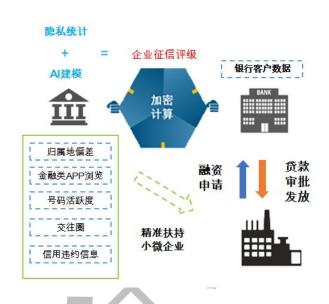


图 3 基于隐私计算的企业风控体系

小微企业信贷风控方面,如图3所示,通过引入企业法人归属地偏差、政企业务欠费情况、号码活跃度、交往圈稳定情况、信用违约信息等运营商侧标签数据,助力银行准确识别小微企业集群背后的复杂关系链条和欺诈风险,构建安全、高效的智慧风控平台,提升银行整体风控水平,助力实现银行小微企信贷业风险识别的精准化、身份核验手段的多样化、提供融资服务的差异化,为银行进行企业信用评价和风险防范提供辅助依据。在数据使用过程中通过采用隐私计算技术进行多方联合统计、联合模型训练及联合预测,保障各方原始数据不出拥有方本地;按用法、用量进行数据定向授权管理等功能保证数

据不会被第三方缓存、转售或二次使用。

2.银行联合精准营销

随着大数据及人工智能的飞速发展,可应用于金融营销的数据维度不断丰富,单一金融机构本身的用户画像已经无法满足精准营销的要求,亟需联合多方机构、企业的数据丰富用户画像,提升营销效果。由于相关法律法规对数据安全、隐私保护等方面的管理日趋严格,各方持有的数据无法跨私域应用,限制了多机构间的数据合作,使多方联合精准营销的应用发展陷入瓶颈。通过隐私计算技术对多方数据进行联合建模,在保障数据安全和数据不出私域的情况下,加强不同金融机构之间、金融机构与其他第三方机构间的数据价值融合,优化营销模型。

金融机构为了更好地服务客户、提升服务质量,会常规性引入外部名单类数据产品进行客群的质量判断和风险判断。考虑到原始数据交互的模式既无法保护用户隐私,也无法避免数据被缓存,近年来已逐步向隐私信息检索(Private information retrieval,简称 PIR)的方式迁移。例如,在提供个性化服务时,为了有效利用客服团队的资源,需要对重点客户进行定向化贴身服务。通过 PIR 的方式,可对运营商提供的 VIP 客群清单在用户隐私受保护模式下进行查询,在不暴露用户个人信息的情况下,完成命中与否的判断。根据查询结果采用差异化服务的方式,合理利用有限客服资源,最大化挖掘、获取客户价值。

如图 4 所示,金融机构在拓展其自身业务过程中,使用隐私计算能力融合多方数据并与其营销能力整合的模式具有显著的效果。例如,

金融机构与流量平台进行数据融合后优化投放客群筛选,提升响应率模型;同时,结合通信运营商数据将风控模型前置,对高净值客群进行提前筛选,并与响应率模型结合,有效提升营销全流程的转化率。



图 4 隐私计算增强下的投放模型与风险模型前置

在银行线上数字化营销方面,借助隐私计算技术,在保障数据交换时数据安全、隐私保护、合法合规的前提下,根据银行目标客户定位,融合银行自有数据和运营商数据,如运营商业务消费等级、网购偏好、用户忠诚度信息、运营商信用评价等,帮助银行构建更加全面的客户画像,快速准确地筛选目标客群,减少无效的营销投放,提高营销精准度,节约大量成本。如图 5 所示。

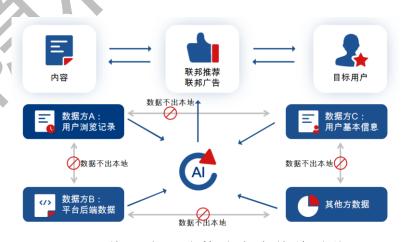


图 5 基于隐私计算的广告推荐系统

通过运营商的用户触达和数据网关等核心能力,可以根据运营商 用户历史活跃数据的支持以及个性化触发能力,智能决策最合适的用 户触达时机和方式,解决用户触达难的问题,如图 6 所示。



图 6基于用户画像筛选广告推荐客户

通过联合筛选、联合建模、联合预测等实现运营商和银行平台 B 端、C 端数据"虚拟打通",从寻客、触达、获客到留存的全链路优化。 在安全融合运营商数据价值后,优化银行的广告投放 ROI,提升客户 留存率,实现精准营销,如图 7 所示。

AI精准营销解决方案全貌



图 7基于隐私计算的精准推荐全流程

在信用卡流失用户挽留方面,可以使用银行内部数据和运营商数据检测某账户所属用户是否有流失风险,并根据用户的行为信息判断用户价值,对流失可能性较大的高价值用户采取一定的挽留措施。银

行侧数据能够反映客户的活跃度发展趋势,而运营商业务消费等级、超前消费意愿、理财意愿、消费活跃程度等数据能够反映客户近期是否仍有信用卡使用意愿。

如图 8 所示,采用隐私计算的方式安全打通银行和通信运营商数据,从而对客户当前情况进行更加精准的分析。使用双方提供的数据进行联合建模,基于联合模型进一步预测用户是否存在流失的风险,并对用户进行评分排序,精准定位出高价值关键客户群体,帮助客户经理制定针对性的挽回及营销策略,预防高价值客户流失,降低银行损失,节约银行开发客户所花费的成本。

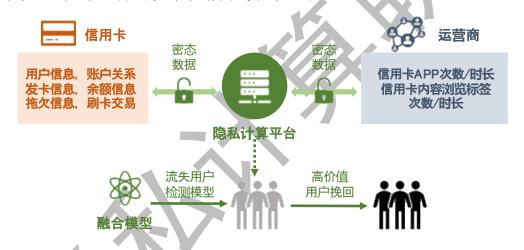


图 8 基于隐私计算的精准推荐全流程

3.保险智慧出险管控

当前,车险保费收入占中国财险份额比例约 50%-60%,是财险绝对主力构成。作为一类金融服务产品,其面临着高赔付、高频率、趋于饱和的业务特性,产品收益的核心体现于出险率和理赔成本的博弈。在商业车险改革进一步深化的行业背景下,进一步加强车险风险管理,运用前沿数据安全融合技术手段,通过数字化、智能化的方式,提高

保险公司车险风控管理综合能力,将是车险风险管理发展的趋势。

通过运用前沿技术手段,依托大数据、AI 机器学习、知识图谱等技术,保险公司可以围绕通信运营商数据、车管局数据及保险自有数据等多方数据共建更加全面的数据智能模式,弥补保险公司自有数据不足的问题。随着用户隐私保护和数据安全的要求和限制愈发严格,通过汇集明文数据进行数据分析和价值挖掘的传统方式面临诸多问题。为化解多方数据价值释放与数据安全保护的矛盾,提升车险风险管理效率,通信运营商、保险公司和车辆管理部门多方合作,采用多方安全计算、联邦学习等技术,针对车险高赔付场景人群、人伤赔付高频场景人群进行风险预警评估。

存量车险用户风控需求主要集中在网约车车主识别、出险高风险 行为人群识别、车险用户画像三项业务场景,借助通信运营商具备的 海量连续性数据、多元化标签数据、丰富大数据处理经验,可助力保 险机构通过前置风险管控的手段,降低高频赔付业务占比。通过隐私 计算平台实现对业务场景执行任务的总体调度及加密后数据的逻辑 运算,各业务参与方通过本地部署的明文计算引擎及数据加解密模块 接入隐私计算平台,实现本地明文数据不出域,各方仅输出密文计算 因子参与业务逻辑运算,如图 9 所示。

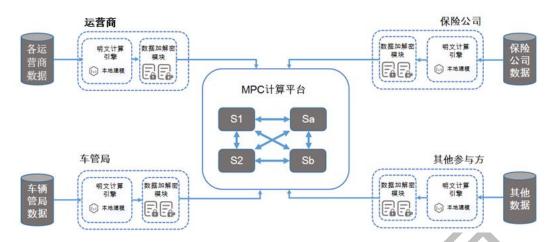


图 9 基于隐私计算的保险智慧出险计算框架

在传统车主风险评估方面,运用保险公司已有的用户历史投保、 出险记录等数据,同时结合运营商自驾游偏好情况、夜间出行偏好、 出行移动范围等级、区域速度等数据,结合车管局相关个人违章信息、 高风险道路位置区域信息、出险检测站等数据,进行联合建模,实现 对车主类型识别、出险概率评估等业务场景,如图 10 所示。

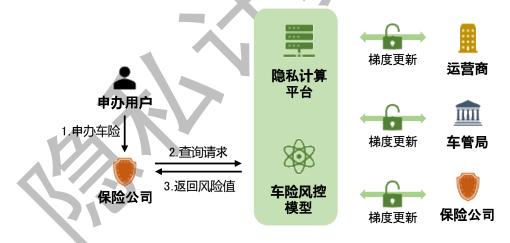


图 10 基于隐私计算的保险智慧出险业务框架

此类项目的实施将帮助保险公司在已趋于饱和的车险风控业务 中实现多维度数据联合建模,既保障了各方数据安全,又增强了建模 所用数据维度,丰富其对投保用户的价值、风险、潜力等多维度评分 矩阵,使保险公司能够在充分了解投保用户的出险概率的前提下,实 现保险行业的业绩及品质提升。

4.证券沉默用户激活

证券沉默用户是指交易活跃度较低且手续费低于一定阈值的客户,这些客户虽为券商注册用户,但并不能为券商带来真实收益。随着股票投资市场的繁荣发展,券商整体客户数量逐年上升,沉默用户的数量也不断累加。部分沉默用户仍具有较高投资意愿,挖掘这部分客户的潜在价值是券商客户运营领域需要解决的关键问题之一。

针对手续费低于一定阈值的证券账户,使用券商内部数据和通信 运营商数据联合检测该账户所属用户是否仍有理财投资意向。券商内 部数据包含用户的资产、收入及交易记录,可以反映用户的交易趋势 和客户价值。运营商数据包含用户投资爱好、财经关注等级、消费活 跃度等数据,可以反映用户的投资意愿。双方数据交叉分析即可得出 该用户是否有激活为活跃客户的潜力和价值。



图 11 基于隐私计算的券商沉默用户激活业务框架

如图 11 所示,采用联邦学习的方法,使用通信运营商的数据补充券商本地数据的不足,从而对客户进行更精准的分析。使用双方的相关数据进行联合建模,检测各沉默用户个体是否仍存在理财投资的

意向,并针对客户价值进行打分排序,精准区分用户圈层,定位出营销投入产出比最高的沉默用户群体,以支撑券商客户运营团队进行精准触达。

(二) 赋能政务

1.电信诈骗识别预警

电信诈骗是一种以非法占有为目的,利用电话、短信、聊天工具等手段,与被害人进行远程接触,通过虚构事实、隐瞒真相或者其他欺骗性手段,骗取他人财物数额较大的犯罪行为。随着现代通信和移动支付技术的迅猛发展,不法分子欺诈手法不断升级,单次欺诈行为贯穿第三方聊天工具、运营商、银行等多个行业及领域。电信诈骗对象不再是对社会资讯相对缺乏的中老年人,而是高度依赖互联网的年轻人。

针对电信反欺诈识别的联邦模型,将运营商的用户静默等级、交际圈稳定程度、运营商业务量变化情况等数据与公安诈骗号码库采用联邦学习技术进行联合建模,实现电信欺诈联合预测。同时,结合已有的欺诈识别策略库,输出电信诈骗名单。通过此名单,能够以电话、短信渠道通知受害人,进而降低财产损失,如图 12 所示。

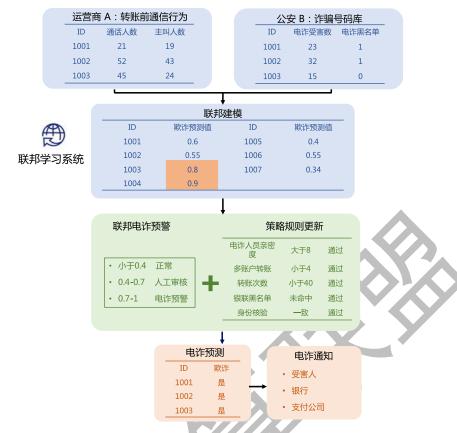


图 12 基于隐私计算的电信反欺诈系统

在该场景中,也可结合区块链技术,为如何确定各参与方的贡献程度、如何进行事后的数据安全审计等问题,提供有效的激励机制与安全审计方案。采用区块链融合联邦学习的技术方案,将所有的交互都通过同态加密和哈希编码技术实现,避免通信运营商和公安方数据的直接传输,仅交互加密的模型中间参数,且交互的动作上传至区块链,加强了数据融通过程中的安全性和可审计性,如图 13 所示。

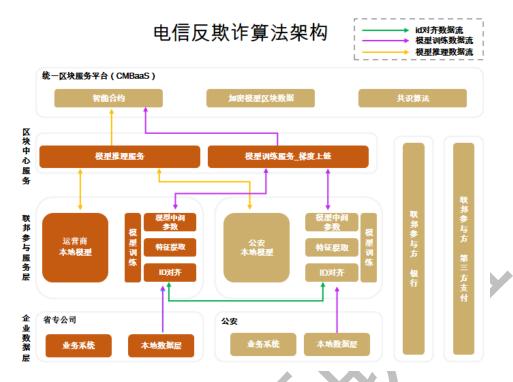


图 13 基于隐私计算的反欺诈算法框架

2.智慧人口流动分析

伴随着当前城市化进程的不断推进,城市规模扩大迅速,城市内 人口数量增长迅猛,各级城市政府单位面临着巨大的管理压力,亟需 推动城市人口治理的创新改革。

以联邦学习为技术核心,打造通信运营商与人社厅的跨域安全联合建模能力,在满足隐私保护、数据安全和法律法规的要求下,基于运营商用户的业务使用数据、位置数据和地方人社厅的标签数据,建立大学生、农民工、企业劳动力等群体态势分析场景,能够为人社厅人口治理和政策优化提供有力支持,如图 14 所示。

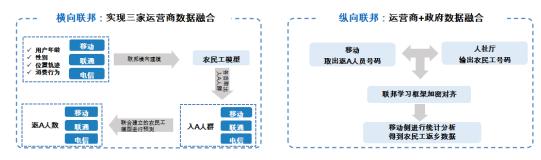


图 14 人社厅+运营商隐私计算建模流程

在此体系下,构建人口流动监测综合解决方案,通过 Web 端以及大屏端展示人口监测应用成果,将通过人工智能手段分析、挖掘出的大学生、农民工等群体流动数据,以图表等可视化形式直观展现并发布,可从宏观和微观多角度了解该区域各群体流动情况,对政策优化起到一定的指导决策作用,如图 15 及 16 所示。

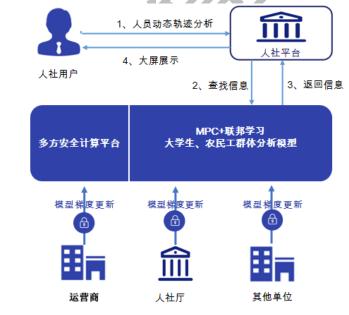


图 15 基于隐私计算的智慧人社全流程



图 16 基于隐私计算的智慧人社分析系统

3.公共安全态势感知

公共安全不仅要保护广大人民群众的生命财产安全,还保卫着地 区稳定和社会发展,但是基层公安日常工作极其繁重,包括:各类人 群管控、受理群众报警求助、强化信息工作、社区安全防护、辖区治 安和消防检查等。向数据要警力,实现对公共安全管理、应急服务的 实时监控分析,辅助主管单位进行安全检查规划和优化工作,用信息 化手段提高公共安全防控能力,避免突发事件的影响,降低生命财产 损失,提高公众服务能力。

在现行的相关法律法规要求下,公安有关部门的数据应严格保密,通信运营商持有的用户样本数据同样不允许泄露。因此,在公共安全领域迫切需要探索和利用新型技术手段,以解决数据流通和数据安全的矛盾,构建数据可信流通环境,提升数字化公共安防服务水平。

依托隐私计算技术构建数据安全流通平台,为公安系统提供模型 上传、运行状态查询和结果集查询等接口。在保障公安隐私数据没有 泄漏风险的前提下,利用多方安全计算技术实现对目标群体聚集地的 统计分析。通过目标群体样本、交往圈、位置特征等信息,完成区域 风险评估模型计算,利用地图上颜色深浅表示事件发生概率高低,使 公安部门更清晰地了解整体情况,达到预防群体性事件发生和管控的 目的。该应用场景能够全面覆盖市级范围预警,在必要时可进一步扩 展至省级范围预警。

在公共安全态势感知场景中,基于安全求交技术,在公安、运营商数据分别不出域的情况下安全融合双侧数据,精准获取地图区域态势感知信息。公安侧以私有云等轻量级方式参与双方交互运算,最终以API接口或图层方式进行结果调用,在底层运用隐私计算新技术的同时保留公安侧原有用户习惯,有效助力社会稳定维护工作,如图 17 所示。

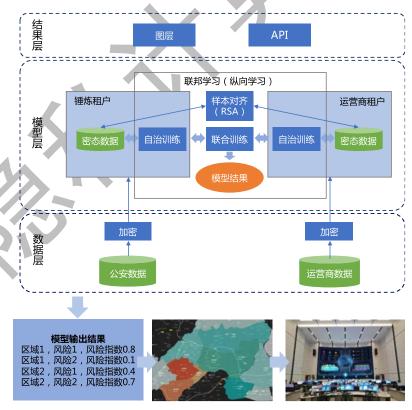


图 17基于隐私计算的公共安全态势感知业务场景

(三) 赋能其他行业

1.医疗精准推荐

健康导航服务平台能够提供预约挂号、排队叫号、报告查询、体检预约、在线药房等各类医疗健康服务。平台方一直在探索互联网运营方式,然而受用户数据量限制,传统营销模式的效果欠佳。运营商拥有用户所处阶段情况、工作时长、夜间活跃程度等维度数据,合理利用通信运营商数据,能够有效加强健康服务平台方在新增场景的存量用户客群推荐、新用户的适用场景推荐等模型的预测效果。因此,亟需通过技术手段在保障双方数据隐私安全的前提下,进行联合模型构建,提升平台推荐模型的准确性。

引入联邦学习技术,在多方本地化部署的基础上,服务端、客户端及协调方通过网络互联进行联合建模,实现原始数据不出库,仅共享数据应用价值,有效解决数据孤岛问题。联合模型充分利用运营商的大数据优势和健康导航平台行业经验,建设健康导航平台弹窗问诊等医疗健康服务功能,实现精准推荐,提升准确率,如图 18 所示。



图 18 基于隐私计算的医疗精准推荐全流程

健康导航服务平台根据自身业务需求发起训练任务,经协调方和通信运营商进行数据对齐,完成后即可与通信运营商交换参数并训练模型。具体步骤如图 19 所示:

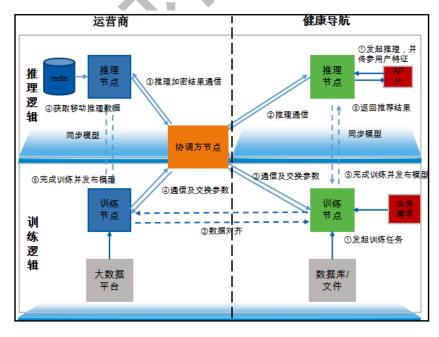


图 19 医疗精准推荐训练流程

在实践过程中,通信运营商与地方卫健委逐步探索联合建立统一 医疗健康服务平台,实现问诊、体检等场景的精准推荐。同时,打造 场景冷启动、用户冷启动等功能模块,以解决对新增场景的用户群推 荐,和新增用户的业务场景推荐的问题,有效提升就医用户满意度。

2.汽车精准营销

自 2018 年以来,中国汽车销量连续 3 年遭遇下跌,行至 2021 年, 汽车运营模式已经"不得不变"。汽车行业在传统模式下曾大获成功, 但到了存量时代,传统获客线索发生退化、单车获客成本上涨、运营 生态悄然改变,相关企业亟须探索数字化营销转型的最优路径,避免 有限的资源被浪费。随着汽车行业的数字化、智能化发展,高质量的 汽车用户使用行为数据成为汽车行业可持续发展的核心要素。但是, 汽车行业存在脱敏数据的隐私与安全性难以保证、数据孤岛、数据监 管等挑战,导致汽车行业数据在多机构之间无法有效完成数据互通共 享,在业务侧表现出潜在客群无体系、对客户业务诉求不明确、客户 服务无追踪、获客效果转化增长低等问题。

针对汽车精准营销的一系列问题,通信运营商依托其在物联网、 云计算、人工智能、大数据等多方面的技术积累,打造汽车产业开放 性的连接生态,联合多方融通数据建立共享的纵向联邦计算平台,最 大化的保障数据安全,最终实现多方安全计算技术与汽车营销场景的 深度融合,为汽车行业提供专业的大数据技术服务支持。通信运营商 与车企共同建立隐私计算平台,确保原始数据仅在本地计算与存储, 最终根据车企客户需求,选择合适渠道进行目标客户触达,如图 20 所示。

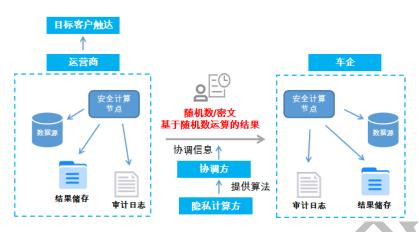


图 20 基于隐私计算的汽车智能营销计算框架

购车意向预测模型能够帮助车企了解潜在购车客户的购买意向 强度、需求量以及预估价位等信息,根据已有线索可进行快速筛选, 圈定高质量种子人群。综合客户消费能力、常驻城市、娱乐偏好、人 生阶段等多维度用户特征数据进行分析,基于多平台、多渠道、多数 据进行联合建模、联合计算,进一步扩展有效训练数据,利用持续迭 代的深度学习算法模型,从海量预选人群中提取高意向目标客群并提 供个性化服务,如图 21 所示。

根据每个人的特征进行广告精准定向

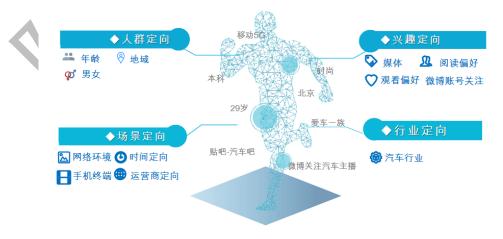


图 21 潜在客户精准画像

运用隐私计算打破数据孤岛,联合建模释放商业价值。一方面, 在挖掘各类车型潜在客户、高价值客户的过程中,能够为企业提供更 优的获客方案。另一方面,也为车企存量用户的保有运营提供策略, 帮助车企进行其客户的分层管理,助力车企产品运营与优化。

四、 通信行业隐私计算创新应用

(一) 智慧城市

随着数字化技术的发展和智能化生活方式的转变,人们对城市的服务能力及服务效率提出了更高要求,促使智慧城市、数字孪生的概念应运而生。通过整合交通、文旅、能源、教育、公共信息等各行业、各领域数据,使城市要素虚拟化、数字化,令网络虚拟空间中的城市与现实中的城市相互映射、实时联系,便捷地实现综合监测、集中管理、治理。

在传统观点中,数据拥有者专注于建设自身的安全体系,隐私保护是一种"被动"的行为,对其他参与方数据安全级别的验证是困难的。然而,智慧城市建设旨在实现政务、交通、产业等城市治理工作数字化,其带来的海量数据中不乏居民信息、征信数据等个人隐私信息或高敏感度的机密数据,其面临的风险包括几个方面:一是,大量隐私数据集中在云平台,使得针对平台发起的攻击格外危险;二是,智慧城市中的许多业务涉及不同政府机构间或企业间,以及跨行业的信息流转,过程中数据可能被相关方沉淀、保存,带来隐私信息泄露的风险;三是,许多场景需要多个机构对数据进行联合挖掘和评估,即各方提供大量的原始数据以得出最终结果,此过程中有原始数据泄露的

风险;四是,流通过程中如果某一方提供了不完整或被篡改的数据,最终结果都可能受到影响,因此保障数据算法的可验证、可度量也至关重要。

隐私计算在智慧城市中的应用将为这种情况带来改观,通过技术手段为各方提供一个安全的数据流通平台,使各方在合作进行数据挖掘、建模或训练时,只能利用到数据价值而看不到数据本身,可以获取计算的结果而无法获取原始数据,并且该安全特性可以通过技术手段进行论证。

例如,在通过治安大数据掌握城市中高危人员、车辆的活动情况以预测和防范公共安全事件的实践中,通过应用可信执行环境与不经意传输技术(oblivious transfer,简称 OT)实现黑盒的数据传输、加工处理。在使用数据时查询请求会被 OT 协议混淆,查询者仅可获取预期的正确数据,且数据提供方无法确定查询者输入的具体信息。

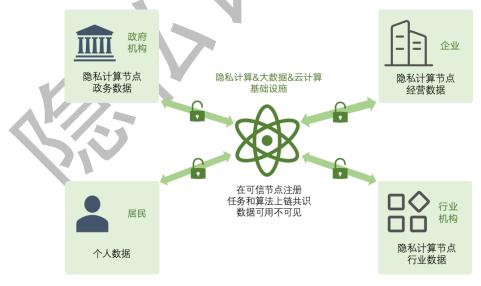


图 22 基于隐私计算的智慧城市系统架构

(二) 云边协同

随着网络部署的逐步完善,5G 网络能力逐步得到体现,其高带宽、低时延等特点能够提升云边协同体系的效率。多接入边缘计算(Multi-access Edge Computing,简称 MEC)作为 5G 关键技术之一,是满足 5G 网络性能和延迟要求,改善客户体验的重要途径之一,能够提供人工智能等应用和服务。

部署大量高性能计算设备的云计算模式能够完成海量数据的计算和存储,但是数据通常需要由终端设备采集上传,数据传输过程会面临传输带宽瓶颈及时延问题。边缘计算既能将本地数据存储在边缘侧,又可将部分计算任务在边缘侧执行,有效缓解了网络传输负载压力,缩短网络通信时延。就近处理的边缘计算作为中心化处理的云计算向用户端的延伸,两者联合构建起云边协同计算体系。边缘端完成本地范围的数据存储和计算工作,云端完成数据的汇聚和全局数据的挖掘分析,进而满足更多的场景需求。

随着《网络安全法》、《数据安全法》、《个人信息保护法》的相继出台,如何安全合规地应用数据成为了当前行业内关注的热点问题。基于隐私计算的云边协同模式为解决这一问题提供了有效技术手段。边缘侧终端收集到的数据通常携带大量个人隐私数据,若直接上传云端存在数据安全和隐私泄露风险,即使只保留在边缘本地,也存在用户隐私数据的被动泄露风险。基于隐私计算的模式能够保证数据不出边缘端本地,由云端协同完成联邦模型的训练,并更新各边缘端本地模型,提供低时延的多样性模型和数据服务。



图 23 基于隐私计算的云边协同系统架构

在云边协同的网络模式中,边缘端通常是数据存储和本地模型训练的一方,云端通常是负责中间结果和模型聚合的一方,在此过程中,隐私计算的三种技术路线:多方安全计算、联邦学习和可信执行环境均能够得到很好的应用。计算过程中不传递原始数据,仅将边缘侧本地计算产生的中间数据通过同态加密、秘密分享等方法传输到云端,有利于保护边缘端的隐私数据。边缘端服务器也可将数据迁移到TEE硬件环境下完成可信计算。

因此,在云边协同的基础上,基于隐私计算的数据与模型应用能够运用在 MEC 框架中,在保障数据安全的前提下实现拓扑结构的智能联接,提供高性能计算以及多样性服务。

目前,业界针对云边协同和隐私计算的技术应用尚处于探索阶段。 在应用隐私计算过程中,数据需要经过各类密码技术处理,会带来较 大的算力开销。并且,在数据传输和计算阶段,相应的密码协议需要嵌入到通信网络协议中,这对网络协议和架构提出了更多的优化要求。随着多源异构网络的融合,万物互联将进一步释放数据驱动力,推动各行业数字化转型发展,云边协同场景应用也将变得更为复杂,终端设备的异构性以及终端数据的异构性、多源性会导致数据安全和隐私保护问题日益严峻,隐私计算在云边协同场景下的需求也将得到更多维度的释放。

(三)算力网络

算力网络是一种根据业务需求,在云、网、边之间按需分配和灵活调度计算资源、存储资源以及网络资源的新型信息基础设施。在算网场景下,算力分布泛在,算力提供者提供算力资源,算力消费者使用算力资源进行数据的计算、存储。

传统模式下,算力消费者将需要计算的数据进行网络传输,传输过程中使用经典加密算法对数据进行保护,算力资源节点对数据进行解密,使用解密后的原始数据计算、存储,数据的所有权发生转移,可能面临无限制的拷贝、复制风险,数据隐私也很难得到有效保护。

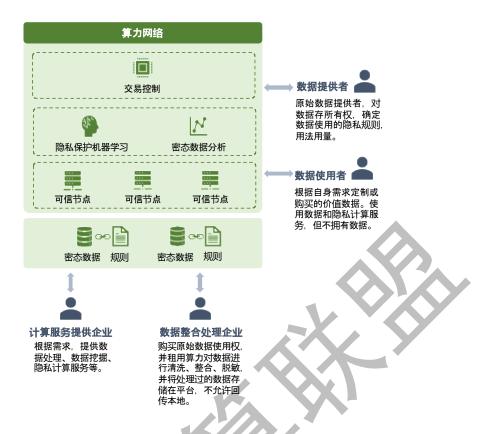


图 24 基于隐私计算的算力网络系统架构

在基于隐私计算的可信算力交易模式下,算力消费者提供的不是原始数据,而是数据的使用权,根据需求确定交易数据的隐私保护规则、交易数据的用法与用量。算力资源节点根据确定的规则采用各类隐私计算技术实现安全的数据汇聚计算,得到的计算结果通过零知识证明等方式验证计算正确性。该模式有效避免了参与方提供数据的拷贝与复制风险,保障隐私数据不被泄露。

结合隐私计算的算力网络交易模式具备以下三点优势:一是原始数据不离开本地,降低数据提供方的安全顾虑;二是数据可用不可见,数据提供方可控制隐私保护规则和数据用法用量,在挖掘数据价值的同时保障数据安全。三是支持丰富的应用场景,包括联合统计、联合查询、联合建模、联合预测等。

五、总结与展望

借助隐私计算技术,能够在保证数据安全的前提下充分发挥通信数据应用价值,助力各行业数字化发展实践,对充分释放数据要素价值和推动社会经济发展具有重大意义。在传统应用上,顺应时代趋势,增强了数据安全保护能力。在创新应用上,为解决新场景落地过程中面临的实质问题提供了较大帮助。

未来,隐私计算仍面临着安全与性能难以兼顾、异构平台壁垒带 来的"数据群岛"问题等诸多挑战有待解决。

在安全与性能平衡方面,隐私计算通信应用中,不乏一些对性能和安全要求较高的场景,如何在保证安全的同时最大化效率和数据价值,成为未来需要解决的关键问题之一。针对隐私计算安全与性能平衡的问题,未来应关注以下几点:

- 通过算法优化、硬件加速等手段,从技术侧实现创新突破,在 固定的安全水平上,优化性能;
- 形成统一的隐私计算安全评价方法,实现安全可验证、可度量;
- 从应用场景出发,结合数据分类分级,制定符合业务场景需求 的安全分级框架。

在跨平台互联互通方面,在通信数据赋能金融、政务等领域的过程中,通信运营商需要与各行业的相关机构进行连接,多平台部署的情况明显,耗费大量资源并提高了使用成本。针对隐私计算互联互通问题,未来应协同多方助力形成互联生态:

- 建立隐私计算互联互通标准化技术规范;
- 除技术攻关,仍需突破适配应用方业务场景,解决业务问题;
- 技术提供方、数据提供方、应用需求方、标准化组织、检测认证机构等行业多方共同探索,构建完善的互联互通生态网络。



参考文献

- [1] 国务院办公厅. 要素市场化配置综合改革试点总体方案[EB/OL]. 2021. http://www.gov.cn/zhengce/content/2022-01/06/content 5666681.htm.
- [2] 国 务 院 . " 十 四 五 " 数 字 经 济 发 展 规 划 [EB/OL]. 2021. http://www.gov.cn/zhengce/content/2022-01/12/content 5667817.htm.
- [3] 中共中央,国务院. 关于加快建设全国统一大市场的意见[EB/OL]. 2022. http://www.gov.cn/gongbao/content/2022/content 5687499.htm.
- [4] 中国隐私计算产业发展报告(2020-2021),国家工业信息安全发展研究中心, 2021.
- [5] 《Gartner 2022 隐私技术成熟度曲线》, Gartner, 2022.
- [6] Shamir A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.
- [7] Blakley G R. Safeguarding cryptographic keys[C]//Managing Requirements Knowledge, International Workshop on. IEEE Computer Society, 1979: 313-313.
- [8] 隐私计算联盟,中国信通院云大所. 隐私计算白皮书(2021年)[R]. 2021.
- [9] Yao A C. Protocols for secure computations[C]//23rd annual symposium on foundations of computer science (sfcs 1982). IEEE, 1982: 160-164.
- [10] Yao A C C. How to generate and exchange secrets[C]//27th Annual Symposium on Foundations of Computer Science (sfcs 1986). IEEE, 1986: 162-167.
- [11] ARM. ARM Security Technology-Building a Secure System using TrustZone Technology. ARM Technical White Paper, 2009.
- [12] 闫树, 袁博, 吕艾临等.《隐私计算——推进数据"可用不可见"的关键技术》 [M]. 电子工业出版社,2022-03-01.
- [13] 魏凯, 闫树, 吕艾临. 数据要素市场化进展综述[J]. 信息通信技术与政策, 2022(08): 59-64.
- [14] Gentry C. Fully homomorphic encryption using ideal lattices[C]//Proceedings of the forty-first annual ACM symposium on Theory of computing. 2009: 169-178.
- [15] Anati I, Gueron S, Johnson S, et al. Innovative technology for CPU based attestation and sealing[C]//Proceedings of the 2nd international workshop on hardware

and architectural support for security and privacy. New York, NY, USA: ACM, 2013, 13(7).

[16] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Artificial intelligence and statistics. PMLR, 2017: 1273-1282.

[17] Yang Q, Liu Y, Chen T, et al. Federated machine learning: Concept and applications[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2019, 10(2): 1-19.

[18] Dash B, Sharma P, Ali A. Federated Learning for Privacy-Preserving: A Review of PII Data Analysis in Fintech[J]. International Journal of Software Engineering & Applications, 2022, 13(4): 1-13.

[19] Gentry C. A fully homomorphic encryption scheme[M]. Stanford university, 2009. [20] 贾轩, 白玉真, 马智华. 隐私计算应用场景综述[J]. 信息通信技术与政策, 2022,48(5):45-52.

[21] 闫树, 吕艾临. 隐私计算发展综述[J]. 信息通信技术与政策, 2021, 47(6): 1.

