云计算环境下数据安全存储与隐私保护技术创新

薛国东[1] 陈向荣[2] 何珩[3] 凌佳生[3]

1 金投健康（杭州）科技有限公司，浙江杭州，310014；

2 浙江移动有限公司，浙江杭州，310014；

3 浙江移动信息产业集成有限公司，浙江杭州，310014；

摘要：随着云计算技术在政务、医疗、金融、教育等领域的广泛应用，数据呈现出高度集中、动态存储和远程访问的趋势，随之而来的数据安全与隐私保护问题日益严峻。传统的数据保护机制已难以适应云计算环境的复杂性和多变性，亟需基于新技术框架的安全策略与隐私管理模式。本文从云计算环境中的数据存储特性出发，深入分析当前存在的安全挑战，系统探讨基于加密、访问控制、可信计算、多方安全计算等核心技术的创新路径，并结合人工智能与区块链等前沿技术探讨其融合发展对数据隐私保护的促进作用。研究发现，构建分层防御体系、动态权限管理与多源数据融合下的隐私保护机制将成为未来技术发展的关键方向。

关键词：云计算；数据安全；隐私保护；加密技术；区块链；访问控制

DOI：10.64216/3080-1508.25.03.045

引言

云计算为数据的存储与处理提供了强大的弹性资源和便利的服务方式，其“即用即取、弹性伸缩、资源池化”的特征使得组织能够显著降低 IT 成本并提升业务效率。然而，云计算环境下的数据安全性问题并未随之消失，反而因其虚拟化、多租户、数据流动性强等特性而愈发复杂。从用户数据遭遇恶意窃取，到云服务商权限滥用，再到数据传输中的中间人攻击，数据安全和隐私保护面临多重挑战。

隐私权作为公民的一项基本权利，在数据驱动社会中逐渐成为数字文明建设的核心议题。如何在保障数据高效流通与共享的同时，实现对数据拥有者意愿的尊重、对敏感信息的保护、对攻击风险的防范，成为技术创新的重要目标。本文将从存储机制、技术路径和融合趋势三个方面系统探讨云计算下的数据安全与隐私保护技术创新路径。

1 云计算环境中的数据存储风险与挑战

1.1 多租户架构下的数据隔离问题

在云计算平台中，多租户架构是提高资源利用率和运营效率的重要设计理念。多个用户共用底层计算资源，表面上通过虚拟化技术如虚拟机（VM）或容器进行逻辑隔离，但在底层物理层面，这种隔离并非绝对安全。攻击者可以利用共享缓存、总线或其他硬件组件进行侧信道攻击（如 Flush+Reload、Prime+Probe 等），间接推断出其他租户的操作行为甚至数据内容。此外，如果云服务提供商在多租户调度中未严格进行安全配置，例如未隔离日志记录通道、共享内存区域等，极有可能造成信息泄漏。此外，在某些场景中，如高频 API 调用和高并发服务访问情况下，虚拟化平台本身可能出现超分配和资源饥饿，从而引发崩溃或越界操作，形成攻击窗口。对抗这一问题需要从底层硬件信任根构建，到虚拟化软件层的权限边界控制，以及租户之间的流量隔离策略等多维度同步优化。

1.2 云端数据管理的可控性缺失

在传统 IT 架构中，数据完全由企业自身服务器管理，从物理存储介质到网络传输路径均处于自控范围。而云计算环境打破了这一控制边界，用户数据被托管在云服务商的基础设施上，用户虽然仍保有数据所有权，但在实际操作中已难以控制数据的存储位置、使用方式及其完整生命周期。特别是跨地域的数据冗余与备份机制，常常使得数据跨境流动，企业难以明确数据是否已被复制、分析或存档。若云平台未提供透明的数据访问日志或审计功能，则用户根本无法知晓其数据是否被合法调用。再者，在某些云服务如 SaaS 系统中，用户甚至无法直接访问其原始数据，只能通过平台提供的接口访问经过处理后的结果，这种“黑箱式服务”加剧了数据可控性缺失所带来的信任危机。监管层面上，欧盟 GDPR、我国《数据安全法》等都强调数据主权和审慎跨境传输的要求，若企业无法掌握数据的边界和命运，则极易面临合规风险和潜在诉讼压力。

1.3 动态环境下的身份验证难题

云计算的开放性与灵活性虽然提升了访问便利性，

但也带来了身份管理的高度复杂化。在实际应用中，一个企业可能存在多个云账号、数百个服务接口和上千个终端用户，每个身份背后都有可能成为潜在攻击入口。传统的静态密码机制在多端接入、移动办公、API 自动化调用等场景下显得过于脆弱，容易遭遇密码重放、社工钓鱼和暴力破解等攻击。尤其是物联网设备和边缘节点，它们往往缺乏强验证能力，成为攻击者突破防线的“软肋”。此外，云环境中用户权限频繁变更、会话持续时间不定，导致难以进行统一、实时的身份管理。若未引入动态多因素认证或基于行为分析的异常检测机制，一旦凭证泄露或身份被冒用，攻击者可能长时间在系统中“潜伏”，并持续窃取数据或破坏业务流程。因此，构建适应性强、身份上下文敏感的认证与访问机制是保障云安全的关键命题。

2 安全存储与隐私保护的核心技术路径

2.1 加密技术的动态演进与应用

加密技术是保护数据最重要的手段，不管是数据在传输还是保存和处理的时候，在云计算的环境下变得更加重要了。像 AES 这样的对称加密，还有 RSA 的非对称加密，虽然它们已经能满足普通的数据保护需要了，但在云计算需要同时保证能用和安全的情况下，新的加密方法就被发明出来了。比如同态加密能够在数据还被加密的时候直接计算，这样就不用先解密再算然后再加密了，处理的时候数据就不会泄露了。属性加密 ABE 不限制钥匙必须对应具体用户了，而是用像’医生’’急诊权限’这样的属性标签来控制谁能看数据，特别适合需要精细分享数据的场合。还有最近的可搜索加密、零知识证明这些技术，已经在医院、政府、银行这些地方用起来了。这些技术不光让加密后的数据更好操作，也让数据安全从原来的单纯保护变成动态管理，最终达到安全使用的目的。

2.2 多因素访问控制机制设计

只用密码验证的方式现在越来越不安全了，因为现在的黑客手段变得很高级。所以现在很多地方都开始用多重认证方法。这个方法就是说要同时检查你知道的东西比如密码、你带着的东西比如手机或者令牌、还有你自己本身的特点比如指纹或者刷脸，这样多重检查让账号更安全。比如同时用指纹解锁再加短信验证码，这样别人就很难盗号了。另外在管理权限的时候，比如公司里用的 RBAC 系统，它给每个职位设定了固定的权限。不过这套办法太死板了，员工调岗后权限没法及时改。后来出现的 PBAC 系统能灵活调整权限规则，只要设置好条件就能自动处理权限。现在最先进的 CBAC 系统就更聪明了，它会看用户的操作习惯、所在位置、用什么设备这些情况，来动态调整权限，比如说只在上班时间和公司电脑上才能进系统。尤其是在远程办公的情况下，这种系统能根据风险级别随时调整权限，这样就能有效防止越权操作。

2.3 可信执行环境与安全容器技术

可信执行环境TEE这个属于硬件级安全的运行机制，它能在 CPU 里专门搞一个安全区域出来，就算操作系统被攻击了，里面存的重要数据也能保持安全和保密。比如说 IntelSGX 就是典型的 TEE 架构，现在已经被用到很多需要高安全级别的地方，像银行密钥管理、数字钱包这样的业务场景，还有智能合约运行这些方面。在云计算的领域里，TEE 可以保证租户的应用代码和敏感数据不被宿主机或者别的租客偷看到。另外现在大家都用容器化和微服务，所以就有了安全容器这种东西。比起以前的虚拟机，容器更轻便部署更快，但隔离效果差了点。所以得配上像 Seccomp、AppArmor 这些安全组件，还有 gVisor、KataContainers 这样的沙箱技术来加强保护。通过这些手段，每个服务组件都能有自己的资源权限设置，遵循最小权限原则，这样就能阻止云环境里的横向攻击和病毒入侵这些风险。

3 前沿技术驱动的数据隐私保护模式创新

3.1 区块链赋能数据操作可追溯

现在数据在云端传播和共享越来越多了，企业和监管部门都很关心怎么才能让数据使用透明监督管理还有责任判定。以前的数据审计主要是靠集中式日志系统来做的，这种容易被修改或者造假日志，比如删除什么的，要是数据被非法访问或者泄露了的话，很难重现整个过程然后追究责任人。不过区块链因为结构是分布式的，再加上链式存储方式，让它对数据无法篡改和能被验证这些特点有特别大的好处。把区块链技术应用到云计算的环境里，不仅可以完整记录数据访问、修改删除这些操作的所有步骤，特别是在很多用户或者不同机构合作的时候，用分布式共识算法还能保证所有节点上的操作记录都是统一不冲突的。

比如在智慧医疗系统里，当病人同意别人看他的病历的时候，这些查看的动作会被立刻写在区块链上，医院、大夫还有保险公司这些都能一起看到这个动作，这样就能保证操作是真的并且马上记录下来，还能随时查得到记录。金融方面来说，像用户查账户、申请借钱、评估风险这些重要操作都会写在链上，这样可以防止公司员工自己乱操作，让管理更符合规定。特别要说的是区块链不只是能做数据检查的平台，还能用智能合约自

动执行和判断能不能看数据的条件，比如说规定只有在周一到周五早上九点到下午五点才能看数据这样的要求直接让系统自己执行，既能遵守规定又能马上反应。这样代码就是规矩，规矩就是法律的设定，让数据管理自己就能运作还可靠，让云端的数据流转更安全、能控制而且看得清楚。

3.2 联邦学习与多方安全计算

在现在这种数据驱动的人工智能模型建立过程里，传统的那种集中式学习方式虽然说效率挺高、部署起来也方便，但对隐私带来的风险现在变得越来越明显。数据都集中保存的话很容易变成黑客攻击的对象，同时还会造成'数据孤岛'这样的情况，这对模型能力提升会有障碍。联邦学习的出现等于是在隐私保护和数据共享之间架起了一个桥梁，它主要就是在不传送原始数据的情况下，通过在本地训练模型然后再把参数更新上传，最后完成整体模型的聚合。这种训练方法特别适合比如医疗啊、金融、教育这些需要联合建模的场景。

国内很多医院可以在不需要交换患者原始数据的情况下，通过联合学习共同训练癌症诊断模型，既让模型的泛用性得到提高，又减少了数据泄漏的可能。和联合学习配合使用的还有安全多方计算（SMC）这种方法，它是把计算任务进行加密后分散处理，让各方在不用暴露自己数据的情况下一起完成任务。比如说多家银行合作开发信用风险模型的时候，每个银行只用参加加密后的运算步骤，不会让别人看到自己的客户数据，这样就能保证数据存在本地是安全的。这种"可以共享信息又不会泄漏隐私"的技术框架，给行业之间数据不通的问题找到了突破方向。不过现在的联合学习还存在像通讯成本过高、不同模型之间差异太大的问题，研究人员在研究用模型蒸馏、压缩参数、同态加密的联邦优化方法来提高效率和实用性。以后把联合学习和多方计算结合起来，就会成为建设下一代"隐私计算平台"的技术基础，让更多需要保密的行业在符合数据规定的情况下做智能应用开发。

3.3 差分隐私的可量化防护机制

在众多隐私保护的方案中，差分隐私(DifferentialPrivacy)因为它在数学上有严格保证还能计算隐私泄漏量的特点，渐渐变成了企业和政府部门主要选用的数据发布手段。它的工作原理就是往原始数据或者查询结果里加上一些设计的噪音，这样攻击者就不能确定某个具体的人的数据是否存在，这样就能保护个人隐私不被发现。DP 最大的好处就是可以调整，通过设定那个叫隐私预算的 ε 值，来平衡隐私保护力度和数据的有用性：ε 要是设得小的话保护效果更好但数据就不太准了，反过来也是这样。现在这种办法被用在了很多地方，比如政府公开的信息、定位服务还有手机 APP 的数据分析等方面。

美国人口普查局从 2020 年开始在发布人口数据的时候用了差分隐私这个方法，主要为了避免别人通过数据反推出来个人身份。苹果公司把这个技术用来分析用户打字的热词，谷歌也在 Chrome 浏览器收集用户行为数据时候用这个技术，进行所谓的“隐形统计”。不过差分隐私也不是完美解决方案，困难的地方在于要根据不同任务调整噪音怎么加进去，而且怎么确定那个 ε 参数设置得合理。现在业内在研究用强化学习、参数自适应这些技术来开发“智能差分隐私系统”，就是说可以根据数据敏感程度、查询次数这些情况，灵活决定怎么加噪声和加多少。另外差分隐私还在和机器学习结合，发展出像带差分隐私的深度学习、推荐系统这些东西，主要是为了在训练模型的时候就减少数据泄露的可能。可以预见的是，差分隐私会从数据保护的最后防线变成智能系统里的基本配置，成为打造可靠 AI 和安全数据管理的关键支撑。

4 结语

云计算作为现代信息社会的技术基石，为数据流通与服务能力提供了前所未有的支持，但与此同时也带来了极高的数据安全与隐私保护挑战。本文围绕云环境下的数据存储风险、安全技术路径及前沿融合技术进行系统探讨，指出多层次防护、多方协作、技术融合将成为数据安全体系演进的方向。在技术创新之外，还需法律法规、行业标准、组织管理等多元合力，共同构建一个安全、可信、开放的云计算生态体系。唯有如此，才能真正实现“数据可用不可见，隐私可控不外泄”的数字安全新愿景。

参考文献

[1]李彦,刘盛.肯尼亚鲁班工坊建设研究——创立云计算与信息安全专业合作模式[J].职业教育研究.2023,(6).DOI:10.3969/j.issn.1672-5727.2023.06.004.

[2]许广彬.云计算环境下的信息安全防护技术研究[J].电子元器件与信息技术.2023,7(7).DOI:10.19772/j.cnki.2096-4455.2023.7.030.

[3]李超宇.基于云计算的网络信息安全技术研究[J].网络安全技术与应用.2023,(11).DOI:10.3969/j.issn.1009-6833.2023.11.031.