

基于隐私集合求交的大模型双向隐私授权

杨帅¹, 韩益亮^{1,2}, 吴旭光¹, 冯浩康¹, 王圆圆¹

(1. 武警工程大学密码工程学院, 陕西 西安 710086; 2. 武警工程大学反恐指挥信息工程教育部重点实验室, 陕西 西安 710086)

摘要: 针对大语言模型多模态接入场景面临的隐私泄露风险, 以及传统基于角色和属性的访问控制模型在用户属性多元、访问规模大时难以实现细粒度隐私管控的问题, 提出一种基于布谷鸟混淆布隆过滤器的隐私集合求交协议, 并构建支持隐私保护的属性访问控制框架。首先, 设计布谷鸟混淆布隆过滤器结构, 通过引入多重哈希链机制降低元素插入时的碰撞概率, 并采用异或操作对过滤器内元素的存储位置进行混淆处理, 进一步增强其抗攻击能力; 其次, 基于该过滤器构建两方隐私集合求交协议, 在标准半诚实模型下实现了单轮通信开销; 最后, 结合基于属性的访问控制模型, 设计具备双向隐私验证能力的访问控制框架, 利用隐私集合求交技术实现访问策略与用户属性信息的双向隐私保护。实验结果表明, 在属性数量为200、并发访问量为10万的场景下, 所提方案的策略验证耗时相较于基于椭圆曲线密码的密文策略属性加密(CP-ABC)方案降低约55%, 显著提升了高并发访问场景下的处理效率, 为大语言模型多模态接入场景提供了兼顾安全与性能的隐私保护方案。

关键词: 隐私集合求交; 属性访问控制; 隐私保护; 大语言模型

中图分类号: TP309

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025062

Bidirectional privacy authorization for large models based on private set intersection

YANG Shuai¹, HAN Yiliang^{1,2}, WU Xuguang¹, FENG Haokang¹, WANG Yuanyuan¹

1. College of Cryptography Engineering, Engineering University of PAP, Xi'an 710086, China

2. Key Laboratory of Counter-Terrorism Command Information Engineering of the Ministry of Education, Engineering University of PAP, Xi'an 710086, China

Abstract: To address the privacy leakage risks in multimodal access scenarios of large language models (LLMs) and the limitations of traditional role-and attribute-based access control models in achieving fine-grained privacy control with diverse user attributes and large-scale access, a private set intersection (PSI) protocol based on cuckoo-garbled Bloom filter (CGBF) was proposed, and an attribute-based access control framework supporting privacy protection was constructed. The CGBF structure was designed: the collision probability during element in-

收稿日期: 2025-04-09; 修回日期: 2025-06-24

通信作者: 韩益亮, hanyil@163.com

基金项目: 陕西省自然科学基金基础研究计划项目(2025JC-YBMS-664)

Foundation Item: The Natural Science Basic Research Program of Shaanxi Province (2025JC-YBMS-664)

引用格式: 杨帅, 韩益亮, 吴旭光, 等. 基于隐私集合求交的大模型双向隐私授权[J]. 网络与信息安全学报, 2025, 11(5): 101-114.

Citation Format: YANG S, HAN Y L, WU X G, et al. Bidirectional privacy authorization for large models based on private set intersection[J]. Chinese Journal of Network and Information Security, 2025, 11(5): 101-114.

section was reduced by introducing a multiple hash chain mechanism, and the storage positions of elements in the filter were garbled through XOR operations, which further enhanced its anti-attack capability. A two-party PSI protocol was constructed based on this filter, achieving single-round communication overhead under the standard semi-honest model. By integrating with the attribute-based access control model, an access control framework with bidirectional privacy verification capability was designed. And the bidirectional privacy protection of access policies and user attribute information was realized using PSI technology. Experimental results show that in scenarios with 200 attributes and 100,000 concurrent accesses, the policy verification time of the proposed scheme is reduced by approximately 55% compared to the CP-ABC scheme based on elliptic curve cryptography, and processing efficiency in high-concurrency access scenarios is significantly improved. This scheme is demonstrated to provide a privacy protection solution that balances security and performance for multimodal access scenarios of LLMs.

Keywords: private set intersection, attribute-based access control, privacy protection, large language model

0 引言

随着大语言模型 (large language model, LLM)^[1] 的快速发展, 以 ChatGPT、DeepSeek 为代表的智能系统已深度融入社会生产生活。当前, LLM 的接入方式呈现多元化趋势, 如图 1 所示, 主要接入方式有: ①个人用户通过 Web 平台直接交互; ②开发者基于智能体平台 (如 AutoGPT、MetaGPT) 构建领域 Agent; ③企业通过 API 接入集成模型能力; ④通过第三方代理 (如云服务提供商) 间接部署实现弹性扩展; ⑤机构采用本地私有化部署保障数据隐私。这种接入方式的多样性在提升人机交互效率的同时, 也带来了差异化的隐私泄露风险。

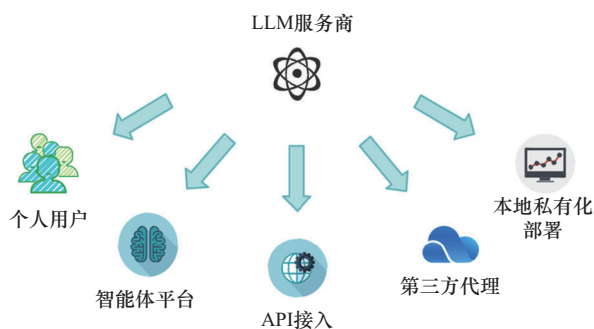


图 1 LLM 的接入方式

Figure 1 Access methods of LLM

不同的用户群体接入需求不同, 不同接入方式面临着特有的安全威胁。个人用户的输入内容可能被平台用于模型微调; 智能体平台中的中间件可能窃取对话上下文; API 密钥泄露易导致服务被滥用; 云环境中的共享部署更容易遭受侧信

道攻击。奇安信资产测绘鹰图平台 2025 年的监测数据显示, 在 8 971 个部署 Ollama 框架的服务器中, 采用云服务提供商部署模式的节点占比高达 63%, 其中 88.9% 的节点未配置任何访问控制机制并直接暴露于公网, 导致攻击者可任意调用服务接口, 甚至非法删除 DeepSeek、Qwen 等模型文件。与传统云服务中以“客户端-服务器”为主的交互方式不同, LLM 支持的多模态交互使攻击面更加复杂。同时, 用户属性在会话过程中可能实时变化, 需要动态更新权限, 这也暴露出当前 LLM 资源访问控制机制在多模态接入场景下存在隐私保护缺陷。

针对上述挑战, 本文聚焦于 LLM 资源访问控制中的隐私保护问题, 提出一种融合密码学技术的创新解决方案。当前主流访问控制模型包括基于角色的访问控制 (role-based access control, RBAC)^[2] 和基于属性的访问控制 (attribute-based access control, ABAC)^[3]。前者通过预定义角色分配权限, 后者依据动态属性判定访问资格。然而, 在 API 调用、智能体平台等开放接入场景下, 传统访问控制机制难以满足大规模用户群体的细粒度隐私保护需求。值得关注的是, 隐私集合求交 (private set intersection, PSI) 技术为该问题提供了新的解决思路。PSI 基于安全多方计算协议, 能够在确保非交集信息不泄露的前提下完成数据比对。进一步地, 将其与属性基加密 (attribute-based encryption, ABE)^[4] 相结合, 可构建支持细粒度权限管理的访问控制机制, 为 LLM 多模态接入环境下的隐私保护提供有效

支撑。

自 Dong 等^[5]首次将布隆过滤器与混淆布隆过滤器引入不经意传输 (oblivious transfer, OT) 协议以来, PSI 领域已取得显著进展。Rindal 等^[6]于 2021 年提出首个基于向量不经意线性估值 (vector oblivious linear evaluation, VOLE) 的 PSI 协议, 显著提升了计算效率。Chen 等^[7]针对数据外包场景设计的 PSI 协议, 在静态半诚实模型下有效保障了数据隐私。然而, 现有研究多聚焦于抗量子密码协议的设计与优化^[8-11], 尚未充分考虑 LLM 系统在动态访问控制、支持海量用户等方面的实际需求。在属性访问控制方向, Ali 等^[12]提出的 AB-PSI 方案通过属性策略约束交集计算, 实现了策略增强型的隐私保护, 但其策略更新机制复杂, 导致较高的计算开销。Cai 等^[13]尝试将 RBAC 与 PSI 结合, 但由于角色一权限映射关系较为固化, 难以灵活适应 LLM 环境下访问需求的动态变化。

针对 LLM 多模态接入场景下用户群体多元、访问规模庞大, 而传统访问控制方案难以兼顾隐私保护与高效验证问题, 本文提出一种融合 PSI 技术与属性访问控制的创新机制。该机制通过策略隐藏实现用户属性与服务器策略的双向隐私验证。此外, 提出了基于布谷鸟混淆布隆过滤器的隐私集合求交 (cuckoo-garbled Bloom filter PSI, CGBF-PSI) 协议, 解决高并发访问场景下的性能瓶颈问题。本文的主要贡献如下。

1) 构建了布谷鸟混淆布隆过滤器 (cuckoo-garbled Bloom filter, CGBF), 通过引入多重哈希链机制, 有效降低了数据存储过程中的哈希碰撞概率。

2) 设计了一种适用于标准半诚实模型的两方 PSI 协议, 实现理论线性复杂度 $O(n)$ 的高效计算, 并仅需单轮通信开销。安全性分析表明, 该协议在半诚实模型下具有可靠的隐私保护能力。

3) 将属性访问控制机制融入 PSI 框架, 通过将策略属性集合与用户密钥直接关联, 简化了传统基于属性加密方案中复杂的访问控制结构, 打破传统访问控制在隐私保护与响应效率间难以平衡的困境, 为 LLM 多模态接入场景提供兼顾安全性与高性能的隐私保护访问控制解决方案。

1 相关工作

PSI 是安全多方计算中的一项关键技术, 旨在解决两个或多个参与方在不泄露各自数据集中非交集信息的前提下, 共同计算出其持有数据的交集。PSI 的核心目标是利用加密算法、哈希函数、不经意传输等密码学技术, 在保证各方数据隐私的前提下, 安全地完成交集运算。在计算过程中, 双方按照既定规则交互经过加密或混淆处理后的数据, 最终仅获取交集结果, 而无法获悉对方集合中的其他任何信息。

在 PSI 的研究历程中, 早期, Kolesnikov 等^[14]提出了基于批处理不经意伪随机函数 (Oblivious PRF) 的 PSI 协议, 为实现大规模数据集的恶意安全计算提供了通信复杂度为线性的高效方案。随后, Ghosh 等^[15]引入了基于不经意线性估值的代数方法, 在两方与多方场景中实现了渐近最优的通信复杂度, 减少了对传统复杂加密技术的依赖。2019 年, Pinkas 等^[16]提出了基于布尔电路并结合 OT 扩展技术的 PSI 协议, 在线性通信复杂度下实现安全计算。同年, Hemenway-Falk 等^[17]基于通用密码学假设构建了适用于复杂协作环境的多方 PSI 协议。Pinkas 等^[18]还提出了轻量级协议 SpOT-Light, 进一步优化了计算与通信开销。Garimella 等^[19]通过改进不经意键值存储 (OKVS) 结构降低了存储复杂度, 为处理更大规模数据集提供了支持。Rindal 等^[6]结合 VOLE 与 OKVS 设计了 Circuit-PSI 协议, 支持对交集元素进行加密后计算并优化了性能。Pinkas 等^[20]基于 PaXoS 框架提出快速恶意安全 PSI 协议, 在缩短运行时间的同时支持硬件加速。2022 年, Raghuraman 等^[21]将改进的 OKVS 与子域 VOLE 技术相结合, 提出了一种在半诚实模型下具有高运行效率和低通信开销的 PSI 协议。总体来看, PSI 技术呈现出“效率不断提升、安全性持续增强、适用场景逐步扩展”的发展趋势。然而, 在面对 LLM 等高并发、低时延的应用场景时, 现有协议在通信轮数与整体开销方面仍难以满足快速响应的实际需求。

为满足更加复杂场景下的安全与隐私需求, PSI 常与其他密码学技术相结合, 以实现更灵活的访问控制机制。例如, 将 PSI 与 ABE 结合, 可

以实现基于属性的细粒度权限管控, 确保只有满足特定属性条件的参与方才能参与 PSI 计算并获取交集结果。Sahai 等^[4]于 2005 年首次提出属性加密概念, 通过属性集合描述用户身份特征, 打破了以往基于身份加密的局限。此后, Zhong 等^[22]提出一种去中心化的 CP-ABE 方案, 借助云存储系统中隐藏的策略实现数据隐私保护与访问控制, 但其计算开销较大。为提升效率与安全性, 周艺华等^[23]基于线性秘密共享访问结构, 构建了云环境中支持策略隐藏的属性基可搜索加密方案。2021 年, Shi 等^[24]提出了结合 PSI 与属性加密的密钥策略。2022 年, Li 等^[25]设计了支持策略隐藏的 CP-ABE 方案, 实现了物联网环境下交集隐私保护的访问控制。2024 年, 刘霞等^[26]针对云环境提出一种策略隐藏且匹配效率较高的 CP-ABE 方案。Ali 等^[12]首次提出基于属性的 PSI 方案 (AB-PSI), 支持数据所有者通过权限策略控制外包数据上的交集计算, 实现细粒度权限管理。2024 年, Cai 等^[13]提出将 RBAC 与 PSI 结合, 实现了物联网系统中的隐私保护授权。然而, 该方案依赖基于 OT 的 PSI 协议, 通信开销较大, 难以应对 LLM 多模态接入场景中数万级并发访问与动态属性验证的效率要求。综上所述, 尽管现有研究在安全性与效率方面已取得显著进展, 但面向 LLM 特有的海量用户、高频交互环境, 仍存在通信轮数多、策略匹配时延大等关键问题, 亟待进一步优化。

2 基础知识

2.1 安全模型

安全多方计算^[27]的安全模型可分为半诚实模型与恶意敌手模型两类理论框架。在半诚实模型中, 攻击者虽遵循协议规范执行操作, 但会收集交互过程中产生的中间状态数据 (包括本地计算记录、传输消息等), 并试图从中推断出其他参与方的隐私输入; 而在恶意敌手模型中, 攻击者可任意偏离协议, 采取如篡改输入、中止协议或伪造消息等主动攻击行为。

本文所提协议基于“理想世界/真实世界”安全范式, 在半诚实模型下达成可证明安全。该范式通过模拟范例完成安全性证明。其核心思想是, 若对于真实世界中任意多项式时间的攻击者, 都能在理想世界中构造一个模拟器, 使其生

成的交互视图与真实执行过程在计算上不可区分, 则称该协议是安全的。具体如下。

半诚实模型: 模拟器通过协议规范与攻击者输入重构交互轨迹, 证明真实协议视图 $\text{View}_{\text{real}}$ 与理想协议视图 $\text{View}_{\text{ideal}}$ 满足 $\text{View}_{\text{real}} \equiv \text{View}_{\text{ideal}}$, 则称协议在半诚实模型下是安全的。

2.2 数字签名

数字签名是一种基于公钥密码体制的核心密码学技术, 用于保障电子数据安全。在签名生成阶段, 签名者使用其私钥对原始消息进行密码学处理: 首先通过哈希函数计算数据的哈希值, 再使用私钥对该哈希值进行加密, 生成数字签名。在验证阶段, 验证者使用签名者的公钥对签名进行解密, 获取哈希值, 同时计算原始消息的哈希值; 通过对比两者是否一致, 可判断数据是否被篡改、签名是否有效, 从而确认签名者的身份。

在本文中, 一个数字签名方案由 Sig 和 Ver 两种算法构成。

Sig 为随机化签名算法, 输入私钥 sk 和消息 m , 输出对应的数字签名 σ 。

Ver 为确定性验证算法, 输入公钥 pk、消息 m 和待验证签名 σ , 若 σ 是 m 的有效签名, 返回 True, 否则返回 False。

2.3 n -out-of- n 加法秘密共享

秘密共享的概念由 Shamir 首次提出, 其初衷是为密码系统构建一种稳健的密钥管理机制^[28]。 n -out-of- n 加法秘密共享是一种基于模加运算的秘密共享方案, 其核心目标是将秘密值 s 分割为 n 个份额, 且仅当 n 个份额齐全时才能重建秘密。该方案的实现步骤如下: 首先, 使用伪随机数生成器随机生成前 $n-1$ 个份额 $s_1, s_2, \dots, s_{n-1} \leftarrow Z_p$; 随后, 利用式(1)计算第 n 个份额。

$$s_n = s - \sum_{i=1}^{n-1} s_i \bmod p \quad (1)$$

在重构秘密时, 需收集全部 n 个份额, 并进行式(2)的模加运算, 即可恢复出原始秘密。

$$s = \sum_{i=1}^n s_i \bmod p \quad (2)$$

由于每个份额 s_i 在 Z_p 上均匀分布, 且缺少任意一个份额时, 剩余份额的联合分布与原始秘密 s 无关, 因此该方案满足完善保密性。若持有的秘密份额少于 n 个, 则无法恢复原始秘密 s 。

一个完整的秘密共享方案由 Genshare 和 Recon 两种算法构成：Genshare 为随机化算法，输入原始秘密 s 和份额数 n ，输出对应的 n 个共享份额；Recon 为确定性算法，输入全部 n 个共享份额，恢复并输出原始秘密。

3 CGBF-PSI 协议

3.1 CGBF 结构设计

基于过滤器结构的 PSI 协议主要依赖高效的哈希函数这一密码学技术来保障安全，因而在轻量级隐私计算中具有不可替代的优势，尤其适用于对通信效率和内存开销敏感的应用场景。然而，此类结构存在固有缺陷，即一定的误判率。以布隆过滤器为例，在存储空间固定的情况下，随着插入元素数量的增加，其误判率会显著上升；若要维持较低的误判率，则需要扩大过滤器容量，从而导致空间利用率下降。

本文提出一种新型的 CGBF 结构，它融合了混淆布隆过滤器与布谷鸟过滤器^[29]的设计思想。其核心机制是通过多列哈希映射与动态冲突消解，实现更高的空间利用率和更低的误判率。具体结构如下。

对于集合 $X = \{x_1, x_2, \dots, x_n\}$ 中的元素，选取 m 个哈希函数，通过 CGBF 构建 $n \times m$ 的二维存储矩阵 M 和位置追踪矩阵 D ，其构造过程如图 2 所示，其中

$$H_{1\dots i}(x) = H_i(H_{i-1}(\dots(H_1(x))))$$

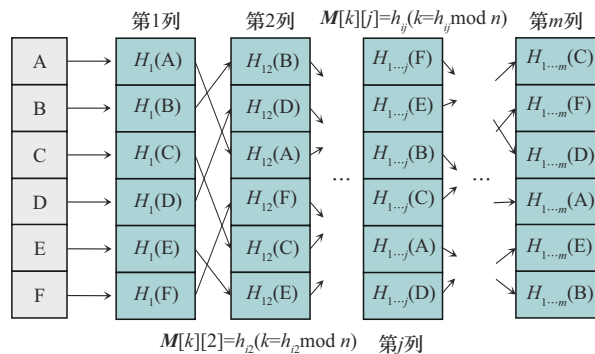


图2 CGBF 存储机制

Figure 2 CGBF storage mechanism

定义1 CGBF 映射规则

1) 首列映射：对于 $\forall x_i \in X$ ，计算 $H_1(x_i)$ 确

定首列存储元素 $M[i][1] = H_1(x_i)$ 。

2) 链式扩展：对于第 j 列 ($j \geq 2$)，计算 $h_{ij} = H_j(H_{j-1}(x_i))$ ，按以下规则定位。

若 $M[k][j]$ 为空，则令

$$M[k][j] = h_{ij}, (k = h_{ij} \bmod n)$$

并记录其位置变化

$$D[k'][j-1] = k, (k' = h_{ij-1} \bmod n)$$

若发生碰撞，递归计算 $k = (k \oplus h_{ij}) \bmod n$ (递归次数小于 n)，当递归次数大于 n 时，则令

$$k = (k \oplus r) \bmod n, (r \in \{0, 1\}^*)$$

直到找到空位，更新 D 链式关系。

定义1中引入的链式扩展机制显著增强了矩阵 M 的混淆性，使攻击者难以从 M 的结构中推断出原始数据的分布特征。此外，CGBF 所采用的冲突消解机制也是一项关键创新。当元素发生位置冲突时，系统将根据预设的跳转规则计算下一个候选位置。若该位置仍被占用，则继续跳转，直到找到可用空间或遍历完所有可能的位置。这一动态调整策略在保障矩阵 M 高效填充的同时，也维护了其混淆特性。

3.2 基于 CGBF 的两方 PSI 协议

在基于 CGBF 的两方 PSI 协议中，协议参与方包括服务提供商 S 和用户 U (如图 3 所示)，共包含系统初始化、数据预处理和安全求交 3 个阶段。

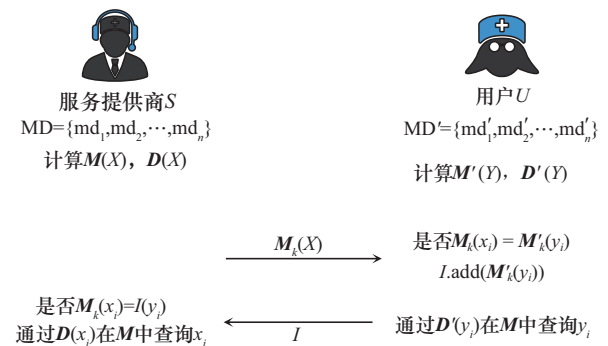


图3 基于 CGBF 的两方 PSI 协议

Figure 3 Two-party PSI protocol based on CGBF

阶段1 系统初始化

1) S 随机选取一些哈希函数 $\{H_1, H_2, \dots, H_k\} \in H$ 并将其发送给 U 。

2) S 、 U 分别生成各自的数字签名密钥对 $(pk_S, sk_S), (pk_U, sk_U)$ 。

阶段 2 数据预处理

1) S 将自己的属性数据集 $MD = \{md_1, md_2, \dots, md_n\}$ 通过 CGBF 插入矩阵 M 中, 同时将元素的位置记录到矩阵 D_S 中。

2) U 将自己的属性数据集 $MD' = \{md'_1, md'_2, \dots, md'_n\}$ 通过 CGBF 插入矩阵 M' 中, 同时将元素的位置记录到 D_U 中, 然后对 M'_k (M'_k 为矩阵 M' 的第 k 列元素) 进行数字签名, 将 $(\text{Sig}(M'_k, sk_U), M'_k)$ 发送给 S 。

阶段 3 安全求交

1) S 接收到 U 发送的 M'_k 后, 首先验证签名, 验证通过, 则将 M'_k 与 M_k 进行对比, 如果 $M_k[i] = M'_k[j]$, $0 \leq i \leq n' - 1, 0 \leq j \leq n - 1$, 那么将 $M'_k[i]$ 添加到交集 I 中。

2) 对 M'_k 中的元素进行遍历后, 可以得到二者经混淆的交集 I 。 S 对 I 执行算法 1 即可在 MD 中找出属性交集。接着, S 对集合 I 进行数字签名, 将 $(\text{Sig}(I, sk_S), I)$ 发送给 U 。

3) U 接收到 I 后, 首先验证签名, 若验证通过, 则将 I 中的元素与 M'_k 中的元素进行对比, 如果 $M'_k[i] = I[j]$, $0 \leq i \leq n' - 1, 0 \leq j \leq \text{len}(I) - 1$, 那么 U 执行算法 1 即可在 MD' 中找出属性交集元素。

算法 1 属性查询

输入 混淆集合 I , 数据矩阵 M , 位置矩阵 D , 属性集 MD

输出 属性交集 MA

```

1)  $MA = \{\}$ 
2) for  $i = 0$  to  $n' - 1$  do
3)  $k = m$ ;
4) if  $M_k[i]$  in  $I$ 
5) for  $z = 0$  to  $n' - 1$  do
6) if  $D^{k-1}[z] = i$ 
7) while  $k > 0$ 
8)  $z = \text{index}(M_{k-1}[z])$ 
9)  $k = k - 1$ 
10)  $MA.add(MD[z])$ 

```

11) break

12) return MA

3.3 安全性证明

3.3.1 安全模型与假设

本协议基于半诚实安全模型, 假设参与方严格遵循协议流程, 但可能通过观察中间结果推断出额外信息。假设服务提供商 S 、用户 U 在一定程度上是可信的, 但可能会试图从各自的操作中获取额外信息。除交集元素外, 任何一方无法获取对方的原始数据。

3.3.2 理想功能定义

定义理想功能 F_{PSI} , 对于输入 S 的数据集 $X = \{x_1, x_2, \dots, x_{|X|}\}$ 和 U 的数据集 $Y = \{y_1, y_2, \dots, y_{|Y|}\}$, 计算交集 $I = X \cap Y$, 并将 I 分别返回给 S 和 U 。

3.3.3 安全性证明

情况 1: S 为诚实且好奇方

首先, 构建模拟器 SIM_S 模拟服务提供商 S 的视图, 给定理想功能 F_{PSI} 输出的交集 I , SIM_S 模拟 S 视图如下。

1) 为确保基础操作环境一致, 模拟器选取与真实协议中完全相同的哈希函数 $H_1, H_2, \dots, H_n$ 。

2) 模拟器生成一个元素由随机值填充的 $\overline{M}_S$ 矩阵。这是因为在真实协议中, S 无法获取 U 的原始数据, 其操作不依赖于 U 的真实数据。

3) 由模拟器模拟 U 的行为, 根据真实协议生成一个结构与真实协议中 M_U 相同但内容为随机值的模拟矩阵 $\overline{M}_U$, 并将 $\overline{M}_U$ 发送给 S , S 接收该模拟矩阵。

4) 模拟器遵循真实协议中规定的交集计算规则, 在模拟数据上计算出交集结果。

$$\tilde{I}: \tilde{I} = \{z \mid \exists i, j, z = \overline{M}_S[i][j] = \overline{M}_U[i][j]\}$$

为证明真实协议中 S 的视图 $\text{View}_{\text{real}}(M_S, M_U, I)$ 和模拟协议中 S 的视图 $\text{View}_{\text{SIM}}(\overline{M}_S, \overline{M}_U, \tilde{I})$ 在计算上不可区分, 本文使用标准的混合假设来证明该定理。本文定义了一系列混合分布 $\text{Hyb}_0, \text{Hyb}_1, \text{Hyb}_2, \dots$, 如果任何两个后续的混合分布在计算上均无法区分, 那么模拟器 SIM 作为一个整体的分布也与实际执行视图相同。具体分析过程如下。

Hyb_0 : 在该混合分布中, 模拟视图与真实视

图完全相同。

Hyb₁: 在该混合分布中, 使用模拟矩阵 $\overline{M}_S$ 替换真实视图中的矩阵 M_S 。由于 S 在实际操作中基于自身数据生成 M_S , 且哈希函数的输出具有随机性和单向性, S 无法根据其自身操作及所接收的信息, 有效区分真实的 M_S 和模拟的 $\overline{M}_S$ 。因此, Hyb₁ 与前一个混合分布 Hyb₀ 在计算上不可区分。

Hyb₂: 在该混合分布中, 使用模拟矩阵 $\overline{M}_U$ 替换真实视图中的矩阵 M_U 。由于哈希函数的随机性与单向性仍然成立, 且 S 无法获取 U 的真实数据, 因此 S 仍无法区分该模拟矩阵与真实矩阵。故 Hyb₂ 与前一个混合分布 Hyb₁ 在计算上不可区分。

Hyb₃: 在该混合分布中, 协议中的交集计算仅依赖于哈希值的对比, 而与原始数据本身无关。基于哈希输出的伪随机特性, S 无法有效分辨其所见的交集结果是由真实数据生成还是由模拟过程产生。故 Hyb₃ 与前一个混合分布 Hyb₂ 在计算上不可区分。

由于上述的混合分布之间都不可区分, 因此 S 的模拟视图和真实视图不可区分, 即:

$$\text{View}_{\text{SIM}}(\overline{M}_S, \overline{M}_U, \bar{I}) \equiv \text{View}_{\text{real}}(M_S, M_U, I)$$

情况2: U 为诚实且好奇方

首先, 构建模拟器 SIM_U 模拟 U 的视图, 给定理想功能函数 F_{PSI} 输出的交集 I , SIM_U 模拟 U 的视图如下。

1) 为确保基础操作环境一致, 模拟器选取与真实协议中完全相同的哈希函数 $h_1, h_2, \dots, h_n$ 。

2) 模拟器生成一个元素随机填充的模拟矩阵 $\overline{M}_U$, 并将其发送给被模拟的 S 。

3) 模拟器接收来自被模拟的 S 返回的模拟交集 $\bar{I}$, $\bar{I}$ 是基于模拟数据按协议规范计算得出。

为证明在真实协议中 U 的视图 $\text{View}_{\text{real}}(M_U, I)$ 与模拟器 SIM_U 生成的视图 $\text{View}_{\text{SIM}}(\overline{M}_U, \bar{I})$ 在计算上不可区分, 本文同样采用混合论证方法进行分析, 具体分析过程如下。

H₀: 在该混合分布中, 模拟器所生成的视图与真实协议执行中的视图完全一致。

H₁: 在该混合分布中, 使用元素随机填充的模拟矩阵 $\overline{M}_U$ 替换真实视图中的矩阵 M_U 。由于 U

无法获取 S 的本地数据, 哈希函数所提供的混淆性保证了矩阵元素的不可关联性, 因此 U 无法区分模拟矩阵 $\overline{M}_U$ 和真实矩阵 M_U 。故 Hyb₁ 与前一个混合分布 Hyb₀ 在计算上不可区分。

Hyb₂: 在该混合分布中, U 所接收的真实交集 I , 被替换为由模拟 S 返回的、基于模拟数据所生成的模拟交集 $\bar{I}$ 。由于计算过程的一致性, U 无法区分真实交集 I 和模拟交集 $\bar{I}$ 。故 Hyb₂ 与前一个混合分布 Hyb₁ 在计算上不可区分。

基于上述分析, 由于相邻混合分布之间均具有计算不可区分性, 根据混合论证原则, 可得出结论: U 的模拟视图和真实视图在计算上不可区分, 即:

$$\text{View}_{\text{SIM}}(\overline{M}_U, \bar{I}) \equiv \text{View}_{\text{real}}(M_U, I)$$

由于对 S 和 U 均能构建出模拟器, 使模拟视图和真实视图在计算上不可区分, 根据半诚实安全模型的严格定义, 证明了基于 CGBF 的两方 PSI 协议在半诚实模型下安全可靠, 即除了数据交集本身的信息外, 双方无法获取对方的其他数据信息, 有效保证了数据的隐私性。

4 基于 CGBF-PSI 的双向隐私验证的属性访问控制框架

4.1 方案框架与核心流程

基于 CGBF-PSI 的双向隐私验证的属性访问控制框架如图4所示, 其中包含认证中心(CA)、服务提供商(S)和用户(U)3个实体。该框架通过以下4个阶段实现策略隐藏与动态更新。

阶段1 系统初始化

1) S 和 U 分别生成签名公私密钥对 (sk_S, pk_S) , (sk_U, pk_U) 。

2) S 随机选取 k 个哈希函数 $\{H_1, H_2, \dots, H_k\} \in H$ 并将其发送给 U 。

阶段2 访问策略生成

(1) 策略编码

LLM 支持多种接入方式 (如智能体平台、API、Web 界面等), 旨在服务于不同的用户群体。不同的用户群体往往具有不同的用户属性。例如, 个人用户的属性会因职业而不同, 企业用

户的属性因行业、规模等特定性质而不同。

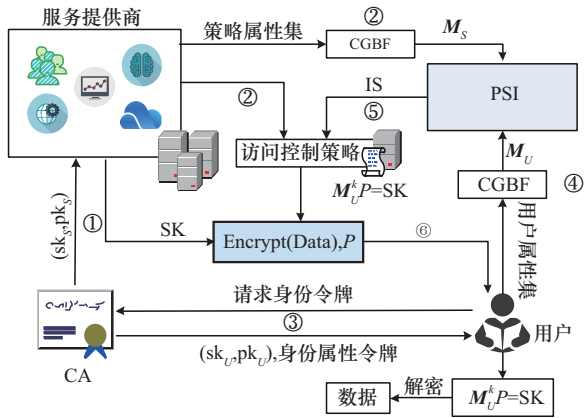


图4 基于 CGBF-PSI 的双向隐私验证的属性访问控制框架
Figure 4 Attribute access control framework for two-way privacy verification based on CGBF-PSI

为应对用户属性与访问需求的多样性,系统须根据实际需求动态生成相应的访问策略。设实际需求集合为 $R = \{r_1, r_2, \dots, r_p\}$, 其中 r_i 代表一种具体的实际需求。基于属性的访问策略集合记为 $DAP = \{dap_1, dap_2, \dots, dap_q\}$, 生成函数表示为 $DAP = G(R)$, 其中 G 是服务提供商根据实际需求生成对应数据访问策略的映射函数, 它基于具体的业务逻辑和安全要求来确定每个 dap_i 的具体规则。访问策略 DAP 全集通过 CGBF 混淆编码为 $n \times k$ 阶矩阵 M_S , 该编码过程表示为:

$$M_S = \text{CGBF}(MA, \{H_1, H_2, \dots, H_k\})$$

(2) 密钥生成及加密

S 为每个策略 $q_i \in DAP$ 生成对应的加密密钥 $K_q = \text{KGen}(q)$, 对服务数据 $d_i \in D$ 进行属性加密 $d'_i = \text{Enc}(d_i, K_q)$, 其中 $q_i \in DAP$ 满足访问条件。

阶段3 请求访问数据

(1) 身份认证与令牌分发

U 向 CA 提交身份信息进行身份验证。通过后, CA 在分发令牌时会依据 LLM 用户群体的多样性及其访问需求, 为不同身份特性的用户分发不同的属性令牌, 从而实现对 LLM 不同服务访问权限的划分。

(2) 属性混淆与签名

U 将属性集混淆编码为 $m \times k$ 阶矩阵 M_U , 并对最后一列 M_U^k 进行签名 $\sigma_U = \text{Sig}(M_U^k, sk_U)$ 。然

后, U 将 $Message_U = (M_U^k, \sigma_U)$ 发送给 S 。

(3) 策略匹配与密钥分割

S 在接收到 U 发送的 $Message_U$ 后, 首先调用验证函数 Ver 对附带的签名进行核验。若 $Ver(Message_U, pk_U) = \text{False}$, 则立即终止流程并拒绝访问; 若 $Ver(Message_U, pk_U) = \text{True}$, 则继续执行后续步骤。

验证通过后, U 须遵循 CGBF-PSI 协议, 计算 UA 与 MA 之间的交集 $IS = UA \cap MA$ 。接下来, S 根据该交集判断用户属性是否满足访问策略, 并据此决定是否授权访问。一旦确认匹配成功, S 随即调用加法秘密共享函数 Genshare , 并据此生成 m 个秘密份额。其表达式为:

$$KS = \text{Genshare}(K) = \{s_1, s_2, \dots, s_m\}$$

然后求解 P , 使得 $M_U^k P = (s_1, s_2, \dots, s_m)$ 。最后对 P 进行数字签名 $\sigma_S = \text{Sig}(P, sk_S)$, 并将 $Message_S = (P, \sigma_S)$ 发送给 U 。

(4) 用户访问并解密数据

U 在接收到 S 发送的 $Message_S$ 后, 验证 S 的签名。若 $Ver(Message_S, pk_S) = \text{True}$, 继续后续操作; 否则, 中止访问。

验证通过后, U 根据 P 求解密钥 K 。具体而言, 首先通过计算 $M_U^k P = (s_1, s_2, \dots, s_m)$ 获得密钥 K 的各个秘密份额, 进而调用秘密重构函数 $\text{Recon}(s_1, s_2, \dots, s_m)$ 恢复出完整的密钥 K 。

最终, U 使用密钥 K 解密其被授权访问的数据 $\text{Dec}(d'_i, K)$, 并使用相关服务。

阶段4 策略动态更新

(1) 策略更新

策略更新包含元素删除和元素添加两种操作。

元素删除: 对于待删除的元素, 系统通过位置矩阵 D 找出对应元素在矩阵 M_S 中的哈希链, 实施删除。

元素添加: 系统首先判断待插入目标行是否已满。若 M_S 不为行全满矩阵, 则直接将待添加元素插入即可; 若 M_S 为行全满矩阵, 则先扩展矩阵行数, 继而完成插入。

S 在策略矩阵更新完毕后, 必须同步执行关

联密钥的更新操作。

(2) 用户属性同步

U 首先依据新生成的策略重新申请对应的属性令牌，继而对该令牌进行必要的编码处理，从而向 S 发起新一轮的访问请求。

4.2 安全性证明

4.2.1 安全模型与假设

本文的安全性证明同样基于半诚实安全模型。在此模型中，假设服务提供商 S 、用户 U 在执行协议时是“半诚实”的，即他们会正确遵循协议流程，但可能会尝试通过记录和分析中间计算结果来推断额外的隐私信息。需要证明数据访问策略的保密性、加密密钥的安全性及数据存储在保密性，确保只有满足特定属性条件的用户才能访问相应数据，并且各方无法获取未授权的信息。

4.2.2 理想功能定义

理想功能 F_{ABAC} 由可信第三方（TTP）执行，其运作流程如下。

1) 输入：接收来自 S 的加密数据 E 和数据访问策略 DAP ，同时接收来自 U 的属性信息 UA 。

2) 判决：根据 DAP 和 UA 进行匹配，判断 U 是否拥有访问权限。

3) 输出：若权限验证通过，将解密密钥 K 返回给 U ，使其能够解密 E 获取数据；否则，拒绝访问。

该理想模型确保了数据的保密性与访问控制决策的正确性。

4.2.3 安全性证明

情况1： S 为诚实且好奇方

在此场景下，需要证明模拟器 SIM_S 能够模拟出真实协议中 S 的视图，且该模拟视图与真实视图在计算上不可区分。首先，构建模拟器 SIM_S 模拟 S 的视图，给定理想功能 F_{ABAC} 输出用户可访问的数据信息（基于访问策略和用户属性）， SIM_S 模拟 S 的视图如下。

1) 为确保基础操作环境一致，模拟器选取与真实协议完全相同的哈希函数 $H_1, H_2, \dots, H_n$ 。

2) 模拟器根据理想功能 F_{ABAC} 输出的可访问性信息，构建一个在访问控制效果上等价于真实策略的模拟策略 $\overline{DAP}$ 。随后，使用该策略所涉及

的属性全集，生成一个结构相同但内容由随机数填充的模拟矩阵 $\overline{M}_S$ 。由于在半诚实模型下 S 无法获知 U 的真实属性，因此模拟操作是可行的。

3) S 接收一个由模拟器 SIM_S 依据协议规范生成的模拟用户属性编码 $\overline{M}_U^k$ 。该编码在结构上与真实编码无异，但内容为随机填充。

4) S 严格遵循真实协议规则，基于模拟的矩阵和属性编码计算模拟交集 $\bar{I}$ 。接着，利用 $\bar{I}$ 和 $\overline{DAP}$ 生成模拟的加密密钥 $\bar{K}$ 和模拟加密数据 $\bar{E}$ （加密过程与真实协议相同，但用模拟的密钥和数据）。

5) S 按照真实协议规则生成模拟的密钥份额 $\overline{KS}$ ，根据协议计算 $\bar{P}$ 并对其进行签名 $\overline{\sigma}_S$ ，将签名后的模拟份额 $(\bar{P}, \overline{\sigma}_S)$ 发送给模拟用户。

为证明真实协议中 S 的视图 $View_{real}(M_S, DAP, M_U^k, E, K, KS, \sigma_S, P)$ 与模拟协议中 SIM_S 的视图 $View_{SIM}(\overline{M}_S, \overline{DAP}, \overline{M}_U^k, \bar{E}, \bar{K}, \overline{KS}, \overline{\sigma}_S, \bar{P})$ 在计算上不可区分，本文采用与第3.3节类似的证明思路，即采用基于标准的混合假设来证明该定理。具体地，本文定义了一系列混合分布 $Hyb_0, Hyb_1, \dots$ ，若任意两个相邻混合分布在计算上不可区分，则可推断模拟器 SIM 所产生的整体分布与真实协议执行的视图在计算上不可区分。具体分析过程如下。

Hyb_0 ：在该混合分布中，模拟视图与真实视图完全一致。

Hyb_1 ：在该混合分布中，模拟器利用理想功能输出的信息构建一个模拟的数据访问策略 $\overline{DAP}$ 。该策略的访问效果与真实协议中的数据访问策略完全相同，但其内部规则由模拟器生成。由于模拟访问策略 $\overline{DAP}$ 与真实访问策略 DAP 的访问效果一致，因此 Hyb_1 与前一个混合分布 Hyb_0 是不可区分的。

Hyb_2 ：在该混合分布中，使用随机生成的模拟矩阵 $\overline{M}_S$ 替换真实视图中的矩阵 M_S 。由于 S 在实际中无法获取用户的真实属性信息，因此替换操作仅依赖于访问策略所定义的属性集合，而与任何用户数据无关。因此，该替换操作不会引入任何可被敌手区分的差异， Hyb_2 与前一个混合分布 Hyb_1 是不可区分的。

Hyb_3 ：在该混合分布中，使用模拟用户属性

编码的向量 $\overline{M}_U^k$ 替换真实用户发送的向量 M_U^k 。 $\overline{M}_U^k$ 由模拟器依据协议规则生成, 结构与真实的 M_U^k 相同, 但数据为随机填充。由于 S 无法获取 U 的真实属性信息, 仅能接收到经过编码的信息 M_U^k , 且所采用的哈希函数具有伪随机性与单向性, 敌手无法区分真实编码与模拟编码。因此 Hyb_3 与前一个混合分布 Hyb_2 是不可区分的。

Hyb₄: 在该混合分布中, 真实的加密密钥 K 被一个由模拟器生成的模拟密钥 $\overline{K}$ 所替代。该模拟密钥是根据随机种子生成的随机序列。 S 无法区分模拟密钥与真实密钥。因此, Hyb_4 与前一个混合分布 Hyb_3 是不可区分的。

Hyb₅: 在该混合分布中, 由真实密钥 K 生成的秘密份额 KS 被替换为由模拟密钥生成的秘密份额 $\overline{\text{KS}}$ 。由于所使用的密钥共享方案具有模拟安全性 (或语义安全性), 模拟器生成的份额与真实份额在计算上不可区分。因此, Hyb_5 与前一个混合分布 Hyb_4 是不可区分的。

Hyb₆: 在该混合分布中, 使用模拟数据的加密结果 $\overline{E}$ 替换真实数据的加密结果 E 。根据加密算法的特性, S 无法区分真实密文与模拟密文。因此, Hyb_6 与前一个混合分布 Hyb_5 是不可区分的。

Hyb₇: 在该混合分布中, 将基于真实数据计算得出的结果 P 替换为基于模拟数据计算得出的结果 $\overline{P}$ 。由于该计算函数是确定性的, 且其所有输入 (数据、密钥等) 在之前的混合步骤中已被不可区分的模拟值替代, 因此其输出结果在计算上也是不可区分的。故 Hyb_7 与前一个混合分布 Hyb_6 是不可区分的。

Hyb₈: 在该混合分布中, 将真实签名 σ_S 替换为由模拟器生成的模拟签名 $\overline{\sigma}_S$ 。该模拟签名并非由真实签名密钥产生, 而是通过模拟算法生成。根据签名方案的存在性不可伪造性 (EUF-CMA) 安全定义, 在未知签名密钥的前提下, 有效的模拟签名与真实签名在计算上是不可区分的。因此, Hyb_8 与前一个混合分布 Hyb_7 是不可区分的。

综上所述, 由于上述混合分布中任意两个相邻的混合分布是不可区分的, 因此 S 的真实视图与模拟视图是不可区分的, 即:

$$\text{View}_{\text{SIM}}(\overline{M}_S, \overline{\text{DAP}}, \overline{M}_U^k, \overline{E}, \overline{K}, \overline{\text{KS}}, \overline{\sigma}_S, \overline{P}) \equiv$$

$$\text{View}_{\text{real}}(M_S, \text{DAP}, M_U^k, E, K, \text{KS}, \sigma_S, P)$$

情况 2: U 为诚实且好奇方

在此场景下, 需要证明模拟器 SIM_U 能够模拟出真实协议中 U 的视图, 且该模拟视图与真实视图在计算上不可区分。首先, 构建模拟器 SIM_U 模拟 U 的视图, 给定理想功能 F_{ABAC} 输出的加密密钥份额, SIM_U 模拟 U 视图如下。

1) 为确保基础操作环境一致, 模拟器选取与真实协议完全相同的哈希函数 $H_1, H_2, \dots, H_n$ 。

2) 模拟器依据协议中定义的属性格式, 生成一个模拟的属性编码向量 (其内容为随机填充), 并为此向量生成一个模拟签名 $\overline{\sigma}_U$ 。随后, 它将此模拟的签名消息 $(\overline{M}_U^k, \overline{\sigma}_U)$ 发送给模拟 S 。

3) 模拟器接收来自模拟 S 的 $\overline{P}$ 和 $\overline{\sigma}_S$ 。

4) 模拟器按照真实协议规定的密钥恢复规则, 计算出模拟密钥 $\overline{K}$ (由于接收到的是模拟份额, 因此计算出的密钥也是模拟的)。接着, 它使用 $\overline{K}$ 对模拟密文 $\overline{E}$ (该密文由模拟器根据理想功能输出的可访问数据信息构建) 进行解密, 得到模拟的解密数据 $\overline{D}$ 。

同样采用混合论证方法证明 U 在真实协议中的视图 $\text{View}_{\text{real}}(M_U^k, P, \sigma_U, \sigma_S, \text{KS}, K, E, D)$ 与其在模拟协议中的视图 $\text{View}_{\text{SIM}}(\overline{M}_U^k, \overline{P}, \overline{\sigma}_U, \overline{\sigma}_S, \overline{\text{KS}}, \overline{K}, \overline{E}, \overline{D})$ 是不可区分的。下文将通过定义一系列混合分布并分析其不可区分性来完成证明。

Hyb₀: 在该混合分布中, 模拟视图与真实视图完全一致。

Hyb₁: 在该混合分布中, 使用由模拟器生成的、内容随机填充的模拟向量 $\overline{M}_U^k$ 替换真实的属性编码向量 M_U^k 。由于协议所使用的哈希函数具有伪随机性与单向性, U 无法区分模拟向量与真实向量, 因此 Hyb_1 与前一个混合分布 Hyb_0 是不可区分的。

Hyb₂: 在该混合分布中, U 接收到来自 S 处的真实向量 P 被替换为一个由模拟器随机生成的模拟向量 $\overline{P}$ 。尽管在真实协议中向量 P 由 S 根据其密钥 M_U^k 和密钥 K 的份额计算得出, 但 U 无法直接获取 S 使用的密钥 K 。因此, 从 U 的视角来看, 模拟向量与真实向量遵循相同的分布, 故

Hyb₂与前一个混合分布Hyb₁是不可区分的。

Hyb₃：在该混合分布中，模拟器遵循真实协议的密钥份额计算过程，但使用之前步骤中引入的模拟向量 $\overline{M}_U^k$ 和 $\overline{P}$ （代替真实输入）来生成模拟的秘密份额 $\overline{KS}$ 。由于该计算过程是确定性的，且其输入（即模拟向量）与真实输入在计算上不可区分，因此其输出份额 $\overline{KS}$ 与真实份额 KS 也是在计算上不可区分的。故Hyb₃与前一个混合分布Hyb₂是不可区分的。

Hyb₄：在该混合分布中，使用模拟份额 $\overline{KS}$ 来恢复密钥 $\overline{K}$ ，将真实视图中恢复出的密钥 K 替换为模拟密钥 $\overline{K}$ 。密钥恢复算法是公开且确定性的。由于 KS 与 $\overline{KS}$ 是不可区分的，且密钥恢复过程一致，因此 U 无法区分模拟密钥 $\overline{K}$ 与真实密钥 K ，故Hyb₄与前一个混合分布Hyb₃是不可区分的。

Hyb₅：在该混合分布中，将真实视图中接收到的密文 E 替换为由理想功能输出的可访问数据加密生成的模拟密文 $\overline{E}$ 。由于加密算法满足语义安全性，在未知密钥 K 的情况下，对不同明文的加密结果在计算上是不可区分的。因 U 无法得知 S 的真实数据，故 U 无法区分模拟密文与真实密文。Hyb₅与前一个混合分布Hyb₄是不可区分的。

Hyb₆：在该混合分布中，将真实密文解密得到的数据 D 替换为使用模拟密钥 $\overline{K}$ 对模拟密文 $\overline{E}$ 进行解密所得的模拟数据 $\overline{D}$ 。由于 K 和 $\overline{K}$ 无法区分且解密过程相同，因此 U 无法区分真实的解密数据 D 和模拟的解密数据 $\overline{D}$ 。故Hyb₆与前一个混合分布Hyb₅是不可区分的。

Hyb₇：在该混合分布中，将 U 在真实协议中生成并验证的数字签名 σ_U 和 σ_S 分别替换为模拟生成的签名 $\overline{\sigma}_U$ 和 $\overline{\sigma}_S$ 。根据签名算法的存在性不可伪造性，模拟器能够在不知晓真实签名密钥的情况下生成在计算上与真实签名不可区分的有效

签名。因此， U 无法区分Hyb₇与Hyb₆。

综上，由于上述混合分布中任意两个相邻的混合分布在计算上是不可区分的，根据不可区分性的传递性，可得出结论： U 在模拟协议中的视图与其在真实协议中的视图在计算上是不可区分的，即：

$$\text{View}_{\text{SIM}}(\overline{M}_U^k, \overline{P}, \overline{\sigma}_U, \overline{\sigma}_S, \overline{KS}, \overline{K}, \overline{E}, \overline{D}) \equiv \text{View}_{\text{real}}(M_U^k, P, \sigma_U, \sigma_S, KS, K, E, D)$$

综上所述，针对协议参与方（ S 与 U ），本文均已成功构建模拟器，并证明其生成的模拟视图与真实视图在计算上是不可区分的。根据安全多方计算中半诚实安全模型的标准定义，能够充分证明本文提出的基于CGBF-PSI的双向隐私验证属性访问控制框架在该模型下是安全的。该协议切实保障了数据访问策略的保密性、加密密钥的安全性及云端数据存储的机密性，确保只有属性满足条件的合法用户才能访问相应服务，且任何一方均无法在此过程中获取未授权的信息。

5 实验和性能评估

5.1 实验环境与方法

本文实验基于Python 3实现，其中哈希函数通过hashlib库实现、加/解密操作通过cryptography库完成。实验硬件配置如下：操作系统为Windows 10，处理器为AMD Ryzen 7 4800H（主频2.90 GHz），内存为16 GB，存储为512 GB NVMe SSD。

5.2 实验设计与结果分析

（1）两方PSI协议性能测试

为评估CGBF结构的性能，本文将其与标准GBF结构及采用同态加密的GBF+结构^[7]展开对比实验。实验配置中，将布隆过滤器的容量设置为1 024，哈希函数数量设置为5。从表1的实验

表1 CGBF、GBF和GBF+元素插入效率对比

Table 1 Comparison of insertion efficiencies for CGBF, GBF, and GBF + elements

方案	数据集大小					哈希函数数量				
	100	200	300	400	500	4	6	8	10	12
GBF	0.002	0.003	0.004	0.005	0.006	0.001	0.002	0.005	0.008	0.012
GBF+	12.57	18.45	22.58	24.41	25.71	10.81	11.68	13.44	15.28	17.11
CGBF	0.008	0.024	0.041	0.071	0.102	0.008	0.009	0.012	0.016	0.024

结果来看, CGBF 的插入效率虽略低于标准 GBF, 但与计算密集的 GBF+相比, 其在插入效率上具有显著优势。

图5直观地展示了数据集大小及哈希函数数量对运行时间的影响。实验结果表明, 运行时间随着数据集的增大而增加, 同时哈希函数数量的增加也会导致运行时间增加。

(2) 访问控制方案性能测试

为评估本文方案在用户属性多样、大规模访问场景下的性能, 本文以文献[30]提出的基于椭圆曲线密码的密文策略属性加密方案为基准, 设计了多组对比实验, 重点探究属性集规模与用户数量对运行时间的影响。

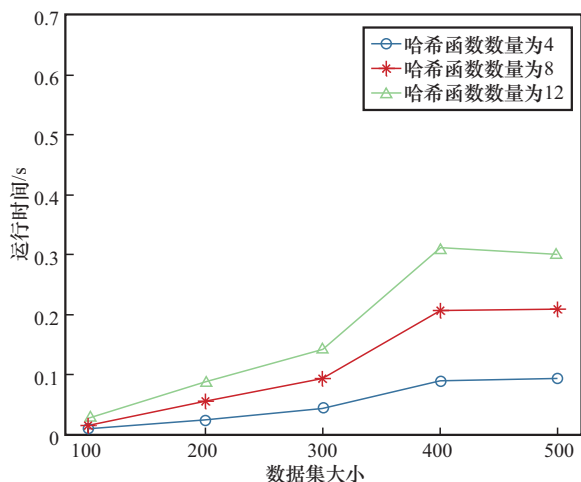


图5 数据集大小及哈希函数的数量对运行时间的影响

Figure 5 The impact of dataset size and the number of hash functions on running time

不同属性集大小和用户数量下的运行时间对比如图6所示。实验结果显示, 随着用户数量的增长, 本文方案的运行时间增长幅度显著低于基准方案。在属性规模扩大时, 基准方案的运行时间虽波动较小, 但其绝对耗时始终高于本文方案。这表明本文方案在应对高并发访问和复杂属性结构时具备更优的效率, 而基准方案在大规模场景下难以满足实际性能需求。

在用户端性能方面, 用户数量和可访问数据量对单个用户运行时间的影响如图7所示。实验结果表明, 在服务器吞吐能力范围内, 用户数量对单个用户运行时间的影响可忽略不计, 而用户可访问数据量与运行时间呈正相关(这主要源于

解密计算开销随数据规模的扩大而增加, 符合理论预期)。

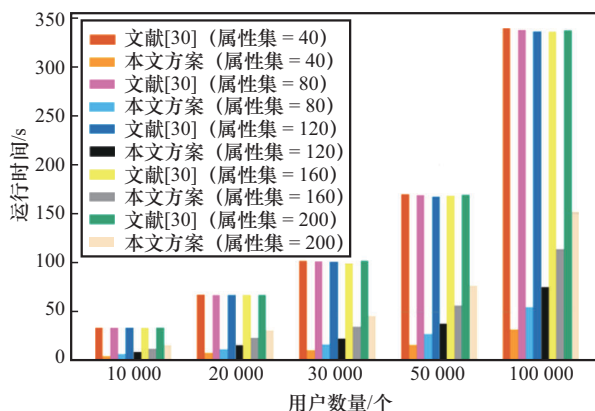


图6 不同属性集大小和用户数量下的运行时间对比

Figure 6 Comparison of running time under different attribute set sizes and number of users

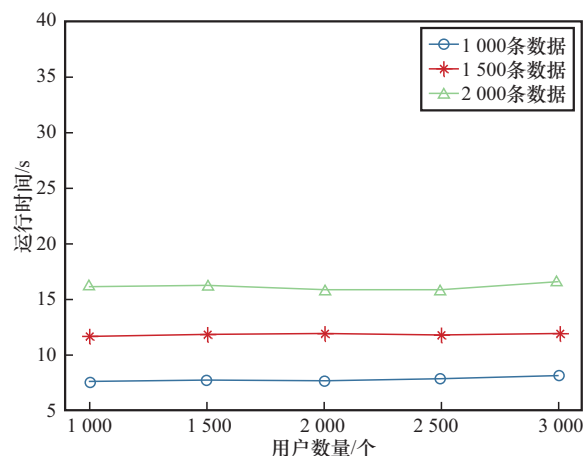


图7 用户数量和可访问数据量对单个用户运行时间的影响

Figure 7 The impact of the number of users and the amount of accessible data on the running time of a single user

综合实验结果表明, 本文方案在大规模访问场景下具有显著效率优势。具体而言, 随着用户数量的增加, 本文方案运行时间的增速仅为基准方案的1/3~1/2; 当属性集规模扩大到120个以上时, 本文方案的运行耗时稳定维持在基准方案的40%~50%, 展现其在复杂属性结构下的高效处理能力。用户端性能进一步验证了本文方案的可扩展性, 在合理负载下基本不受用户规模影响, 而可访问数据量与解密时间的正相关关系符合预期, 反映了数据处理规模对计算开销的合理影响。

综上所述, 本文方案将PSI技术与属性访问

控制机制相结合，在保障安全性的前提下，显著提升了处理大规模用户访问和复杂属性集合的效率。实验数据表明，本文方案在运行时间增速与绝对耗时上均优于传统基准方案，证明其更适用于需要支持海量用户并发接入与复杂策略匹配的应用场景。

6 结束语

针对 LLM 多模态接入场景下面临的用户属性多样、访问规模庞大所带来的隐私保护挑战，本文提出一种基于 CGBF-PSI 协议并支持双向隐私验证的属性访问控制框架。该框架的核心优势在于采用多重哈希链与异或混淆机制优化数据结构，在降低布隆过滤器碰撞概率、增强抗攻击能力的同时实现了高效的隐私集合求交计算；此外，通过策略隐藏机制构建了用户属性与服务器策略之间的双向隐私验证通道，避免了传统属性加密方案的复杂结构，从而显著提升了大规模场景下的策略匹配效率与系统可扩展性。然而，本文仍存在若干不足之处。首先，当前方案基于半诚实安全模型，缺乏对恶意敌手的有效防御机制；其次，在动态策略更新方面，尚未引入轻量级增量更新机制，高频更新场景下策略重构开销可能会影响系统的实时性；再者，LLM 的训练与推理过程涉及大量数据的频繁交互，由此带来的安全问题较复杂，本文方案在此类场景下的整体保护机制尚不完善；最后，在超大规模属性集合与百万级并发访问条件下，内存占用与线程同步开销仍存在进一步优化的空间。

未来的研究工作将重点围绕以下方向展开：增强方案在恶意敌手模型下的安全性，设计支持动态策略的轻量级增量更新机制，并系统考虑 LLM 全流程中的数据访问安全问题，以推动本文方案在实际工程中的应用落地，助力 LLM 生态的安全构建。

参考文献：

- [1] DAS B C, AMINI M H, WU Y Z. Security and privacy challenges of large language models: a survey[J]. ACM Computing Surveys, 2025, 57(6): 1-39.
- [2] SANDHU R S, COYNE E J, FEINSTEIN H L, et al. Role-based access control models[J]. Computer, 1996, 29(2): 38-47.
- [3] HU V C, KUHN D R, FERRAILOLO D F, et al. Attribute-based access control[J]. Computer, 2015, 48(2): 85-88.
- [4] SAHAI A, WATERS B. Fuzzy identity-based encryption[C]//Advances in Cryptology-EUROCRYPT 2005. Berlin: Springer, 2005: 457-473.
- [5] DONG C Y, CHEN L Q, WEN Z K. When private set intersection meets big data: an efficient and scalable protocol[C]//Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security - CCS '13. New York: ACM, 2013: 789-800.
- [6] RINDAL P, SCHOPPMANN P. VOLE-PSI: fast OPRF and circuit-PSI from vector-OLE[C]//Advances in Cryptology - EUROCRYPT 2021. Berlin: Springer, 2021: 901-930.
- [7] CHEN Y C, HUANG K C. JEDI: joint and effective privacy preserving outsourced set intersection and data integration protocols[J]. IEEE Transactions on Information Forensics and Security, 2023, 18: 4504-4514.
- [8] LIU B, RUAN O, SHI R H, et al. Quantum private set intersection cardinality based on bloom filter[J]. Scientific Reports, 2021, 11: 17332.
- [9] HUANG X, ZHANG W F, ZHANG S B. Quantum multi-party private set intersection using single photons[J]. Physica A: Statistical Mechanics and Its Applications, 2024, 649: 129974.
- [10] YE C Q, LI J, YE T, et al. Quantum scheme for private set intersection and union cardinality based on quantum homomorphic encryption[J]. arXiv preprint, arXiv.2412.01032, 2024.
- [11] HOU M, WU Y, ZHANG S B. Quantum private set intersection scheme based on bell states[J]. Axioms, 2025, 14(2): 120.
- [12] ALI M, MOHAJERI J, SADEGHI M R, et al. Attribute-based fine-grained access control for outsourced private set intersection computation[J]. Information Sciences, 2020, 536: 222-243.
- [13] CAI R S, CHEN L X, ZHU Y Z. Using private set intersection to achieve privacy-preserving authorization for IoT systems[J]. Journal of Information Security and Applications, 2024, 83: 103759.
- [14] KOLESNIKOV V, KUMARESAN R, ROSULEK M, et al. Efficient batched oblivious PRF with applications to private set intersection[C]//Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2016: 818-829.
- [15] GHOSH S, NILGES T. An algebraic approach to maliciously secure private set intersection[M]//Advances in Cryptology - EUROCRYPT 2019. Cham: Springer International Publishing, 2019: 154-185.
- [16] PINKAS B, SCHNEIDER T, TKACHENKO O, et al. Efficient circuit-based PSI with linear communication[C]//Proceedings of the 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt. Berlin: Springer, 2019: 122-153.
- [17] HEMENWAY-FALK B, NOBLE D, OSTROVSKY R. Private set intersection with linear communication from general assumptions[C]//Proceedings of the 18th ACM Workshop on Privacy in the Electronic Society. New York: ACM, 2019: 14-25.
- [18] PINKAS B, ROSULEK M, TRIEU N, et al. SpOT-light: lightweight private set intersection from sparse OT extension[M]//Ad-

vances in Cryptology - CRYPTO 2019. Cham: Springer International Publishing, 2019: 401-431.

- [19] GARIMELLA G, PINKAS B, ROSULEK M, et al. Oblivious key-value stores and amplification for private set intersection[M]//Advances in Cryptology-CRYPTO 2021. Cham: Springer International Publishing, 2021: 395-425.
- [20] PINKAS B, ROSULEK M, TRIEU N, et al. PSI from PaXoS: fast, malicious private set intersection[M]//Advances in Cryptology - EUROCRYPT 2020. Cham: Springer International Publishing, 2020: 739-767.
- [21] RAGHURAMAN S, RINDAL P. Blazing fast PSI from improved OKVS and subfield VOLE[C]//Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2022: 2505-2517.
- [22] ZHONG H, ZHU W L, XU Y, et al. Multi-authority attribute-based encryption access control scheme with policy hidden for cloud storage[J]. Soft Computing, 2018, 22(1): 243-251.
- [23] 周艺华, 扈新宇, 李美奇, 等. 云环境下基于属性策略隐藏的可搜索加密方案[J]. 网络与信息安全学报, 2022, 8(2): 112-121.
- ZHOU Y H, HU X Y, LI M Q, et al. Searchable encryption scheme based on attribute policy hiding in a cloud environment[J]. Chinese Journal of Network and Information Security, 2022, 8(2): 112-121.
- [24] SHI Y F, QIU S. Delegated key-policy attribute-based set intersection over outsourced encrypted data sets for CloudIoT[J]. Security and Communication Networks, 2021, 1: 5595243.
- [25] LI J G, ZHANG Y C, NING J T, et al. Attribute based encryption with privacy protection and accountability for CloudIoT[J]. IEEE Transactions on Cloud Computing, 2022, 10(2): 762-773.
- [26] 刘霞, 王馨族, 张涛, 等. 支持访问策略部分隐藏的CP-ABE方案[J]. 通信学报, 2024, 45(10): 180-190.
- LIU X, WANG X Z, ZHANG T, et al. CP-ABE scheme supporting partially hidden access policy[J]. Journal on Communications, 2024, 45(10): 180-190.
- [27] DARBY M L, NIKOLAOU M. MPC: Current practice and challenges[J]. Control Engineering Practice, 2012, 20(4): 328-342.
- [28] SHAMIR A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.
- [29] FAN B, ANDERSEN D G, KAMINSKY M, et al. Cuckoo filter: practically better than bloom[C]//Proceedings of the 10th ACM International on Conference on Emerging Networking Experiments and Technologies. New York: ACM, 2014: 75-88.
- [30] CHENG R, WU K H, SU Y L, et al. An efficient ECC-based CP-ABE scheme for power IoT[J]. Processes, 2021, 9(7): 1176.

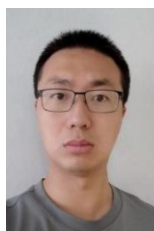
[作者简介]



杨帅 (2002-), 男, 甘肃平凉人, 武警工程大学硕士生, 主要研究方向为隐私保护与密码学。



韩益亮 (1977-), 男, 甘肃会宁人, 武警工程大学教授, 主要研究方向为密码学与机密计算。



吴旭光 (1986-), 男, 河南商丘人, 武警工程大学讲师, 主要研究方向为区块链与网络安全。



冯浩康 (2002-), 男, 河北邯郸人, 武警工程大学硕士生, 主要研究方向为密码学与信息安全。



王圆圆 (2001-), 女, 江西上饶人, 武警工程大学硕士生, 主要研究方向为密码学与数据聚合。