软件学报 ISSN 1000-9825, CODEN RUXUEW
2026,37(2):875−893 [doi: 10.13328/j.cnki.jos.007426] [CSTR: 32375.14.jos.007426]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案*

李 鲍, 周福才, 王 强, 冯 达

(东北大学 软件学院, 辽宁 沈阳 110169)

通信作者: 周福才, E-mail: fczhou@mail.neu.edu.cn

摘 要: 随着物联网的快速发展, 越来越多智能终端设备采集大量患者的医疗数据进行辅助医疗, 具有十分重要的医疗研究价值. 然而, 这些医疗数据通常涉及患者的敏感信息, 且医疗数据在聚合和传输过程中可能面临数据篡改和未经授权访问等安全问题. 为了解决上述安全与隐私问题, 同时支持医疗数据的细粒度的聚合统计分析, 提出了基于 TEE (trusted execution environment) 安全高效细粒度统计分析与可验证数据聚合方案. 该方案对 m 与 m^2 双消息类型 BGN 同态加密算法进行了改进, 并结合了数字签名等技术, 确保了医疗数据的机密性和完整性. 采用了一种可验证的聚合签名算法, 实现了医疗密文数据的批量验证, 降低了认证成本. 通过将医疗密文数据复杂的统计分析过程转换成为 TEE 内的统计分析过程, 提高了医疗数据的统计分析的效率, 同时也降低了计算代价. 采用边缘服务器对研究中心进行授权访问的机制, 实现了医疗数据的细粒度统计分析. 在性能分析方面, 该方案在统计分析侧和数据拥有者侧的计算开销方面具有明显优势.

关键词: 聚合签名; BGN 同态加密; 可信执行环境 (TEE); 统计分析; Intel SGX; 访问控制

中图法分类号: TP309

中文引用格式: 李鲍, 周福才, 王强, 冯达. 基于TEE安全高效的细粒度统计分析与可验证数据聚合方案. 软件学报, 2026, 37(2): 875–893. http://www.jos.org.cn/1000-9825/7426.htm

英文引用格式: Li B, Zhou FC, Wang Q, Feng D. Secure and Efficient Fine-grained Statistical Analysis and Verifiable Data Aggregation Scheme Based on TEE. Ruan Jian Xue Bao/Journal of Software, 2026, 37(2): 875–893 (in Chinese). http://www.jos.org.cn/1000-9825/7426.htm

Secure and Efficient Fine-grained Statistical Analysis and Verifiable Data Aggregation Scheme Based on TEE

LI Bao, ZHOU Fu-Cai, WANG Qiang, FENG Da

(Software College, Northeastern University, Shenyang 110169, China)

Abstract: With the rapid development of the Internet of Things (IoT), a growing number of smart terminal devices collect large volumes of patient medical data to support healthcare applications, offering considerable value for medical research. However, such data typically involve sensitive patient information and may face security risks such as tampering and unauthorized access during aggregation and transmission. To address these security and privacy concerns while enabling fine-grained statistical analysis, this study proposes a secure and efficient statistical analysis and verifiable data aggregation scheme based on trusted execution environments (TEE). The proposed scheme improves the m and m^2 dual-message BGN homomorphic encryption algorithm and integrates digital signatures to ensure data confidentiality and integrity. A verifiable aggregate signature algorithm is introduced to enable batch validation of encrypted data, thus reducing authentication overhead. By shifting the complex statistical analysis of ciphertext data into the TEE, the scheme enhances computational efficiency while reducing processing costs. Moreover, fine-grained statistical analysis is achieved through an access control mechanism based on edge servers that authorize research center access. Performance evaluations indicate that the proposed scheme significantly reduces computational overhead on both the statistical analysis and data owner sides.

* 基金项目: 国家自然科学基金 (62072090, 62202090, 62173101); 中央高校基本科研业务费专项资金 (N2417006); 辽宁省博士科研基金 (2022-BS-077); 辽宁网络安全执法协同创新中心课题 (XTCX2024-015)
收稿时间: 2024-01-26; 修改时间: 2024-11-19; 采用时间: 2025-03-10; jos 在线出版时间: 2025-10-29
CNKI 网络首发时间: 2025-10-31

Key words: aggregate signature; BGN homomorphic encryption; trusted execution environment (TEE); statistical analysis; Intel SGX; access control

云计算[1–4]的出现极大地改变了传统的计算模式, 可以为用户提供更为便捷的计算和存储服务. 同时, 随着物联网技术的快速发展, 越来越多的智能设备接入物联网. 例如, 智慧医疗、智能交通及智慧政务[5–7]等应用的兴起, 万物互联使得网络数据呈现爆炸式增长. 然而这些海量数据涉及大量敏感与隐私信息, 这就需要采用更加安全且高效的方式进行处理与分析. 鉴于传统云计算架构已难以满足对海量医疗数据的实时统计与分析需求, 因此亟需引入边缘计算服务器[5–7], 以协同云端系统实现对医疗数据的实时处理与分析. 虽然, 边缘服务器可分担云服务器的部分计算任务来提高数据处理的效率, 但在聚合分析和数据隐私方面依然不乐观. 特别地, 在智慧医疗领域中, 智能穿戴设备采集的医疗数据往往都涉及患者的敏感信息. 因此, 数据拥有者通常采用智能穿戴设备上的加密技术来保障其医疗数据的机密性和隐私性[8]. 然而, 医疗数据在传输与存储阶段可能面临着通信信道被外部攻击者窃取、篡改[9]以及未授权访问[10]等安全威胁, 若医疗研究中心采用篡改后的医疗数据进行临床分析或医疗研究, 可能危及患者生命安全健康或造成财产重大损失. 此外, 也需要采用访问控制机制实现对医疗数据细粒度的统计分析与安全管控. 同时, 还需要采取数字签名[11–16]与哈希函数来确保密文数据在存储和传输过程中的完整性.

事实上, 加密后的医疗数据会丧失不同程度的可用性, 进而导致了研究中心很难对其进行更为精准的统计分析. 近年来, 诸多研究者利用同态加密[17–19]算法加密后的密文数据仍具有加法和乘法的特性, 提出了许多密文数据聚合方案[20–26], 实现密文数据安全快速的聚合分析. 然而, 这些聚合方案[24,25]大多数都需要研究中心利用私钥直接对聚合密文进行解密得到聚合的结果, 再利用聚合结果进行统计分析, 这样大大增加了研究中心的计算代价与通信代价. 文献 [20–26] 提出的数据聚合协议可在合理时限开销内汇总海量数据, 而无需对单个密文进行逐一解密, 从而实现医疗数据隐私性与机密性的安全保障. 但这些方案[20–26]得到的聚合值仅能满足粗粒度统计分析需求, 无法支持更为细粒度的统计分析. 因此, 亟需设计一种能够同时满足医疗数据的机密性、完整性和用户隐私性等安全需求, 又能实现细粒度统计分析的方案.

综上, 本文针对医疗数据设计并提出了一种基于 TEE (trusted execution environment) 安全高效的细粒度统计分析与可验证数据聚合方案. 该方案针对双消息类型 (m 和 m^2) BGN 同态加密算法进行了改进, 并结合一种基于身份的聚合签名算法, 从而实现医疗数据机密性与完整性双重保障, 同时也实现了医疗密文数据的批量验证. 该方案将边缘计算集成到具有 TEE 的云计算架构, 实现了医疗密文数据统计分析向医疗明文数据统计分析的转化, 提升了医疗数据统计分析的效率. 最后, 采用访问控制手段实现了 TEE 内部的细粒度统计分析, 不仅防止了医疗数据的未经授权访问, 也降低了研究中心的计算代价. 主要贡献如下.

(1) 机密性: 提出了一种优化的双消息类型 BGN 同态加密算法 (optimized double message type BGN homomorphic encryption algorithm, ODMT-BGN), 该算法不仅保留了对原有两种格式 (m 和 m^2) 数据的加密特性, 还在同等安全参数下提供了更高的安全性.

(2) 批量验证: 为了实现对医疗密文数据的批量验证, 本方案采用身份绑定的聚合签名算法, 实现了医疗密文数据在不同存储和传输阶段的批量验证, 进而防止医疗密文数据在存储和传输过程中被篡改或替换, 同时也降低了认证成本.

(3) 高效性: 考虑到医疗密文数据统计分析效率低以及 TEE 的特性, 将原本的密文数据统计分析过程移植到云服务器的 TEE 内, 使其变成明文数据统计分析. 此设计不仅提高统计分析的计算效率, 还降低了研究中心的计算代价.

(4) 细粒度的访问控制: 为了实现医疗密文数据的安全管控, 采用边缘服务器对研究中心进行授权访问的机制, 实现了医疗密文数据的安全管控, 进而通过聚合不同边缘服务器的授权密钥实现细粒度的统计分析.

本文第 1 节介绍了可信执行环境 TEE (例如 Intel SGX)、可验证数据聚合和安全数据共享的相关工作. 第 2 节介绍了本文所需基础知识, 包括双线性映射、决策双线性 Diffie-Hellman 假设、计算性 Diffie-Hellman 假设、哈希函数以及新型 BGN 同态加密. 第 3 节给出了基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案,

包括 ODMT-BGN 算法的设计、系统模型、威胁模型与方案详细描述. 第 4 节给出了方案的正确性证明和安全性分析. 第 5 节是该方案的实验分析. 第 6 节总结全文.

1 相关工作

1.1 Intel SGX

可信执行环境 (TEE) 是一种安全的计算环境, 确保其内部敏感数据和代码的安全性和隐私性. TEE 通常是由硬件和软件组成的, 可以在普通的操作系统上提供一种安全的运行环境, 使得恶意软件无法获取或篡改 TEE 中的数据和代码. 例如, Intel SGX (software guard extension).

英特尔软件防护扩展 (Intel SGX/TDX)[27–29]于 2015 年首次在第 6 代 Intel Core 处理器上实现, 是一组添加在 Intel 架构处理器内一组 x86 内存访问指令和机制. 被保护的内存区域称飞地 (enclave) 驻留在应用程序的地址空间, 旨在为应用程序代码和数据提供基于硬件的内存加密和隔离, 保证了代码和数据机密性与完整性. 此外, Intel SGX 提供了内存隔离、远程认证以及数据封装等功能. 如图 1 所示.

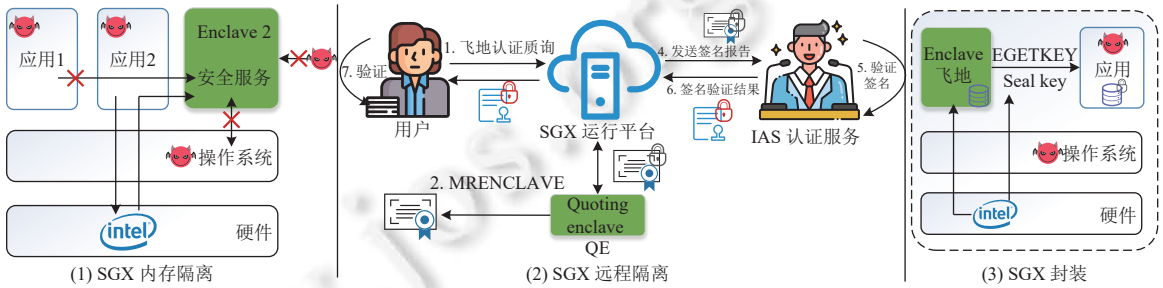


图 1 可信执行环境 Intel SGX

(1) 内存隔离: 飞地的内存被映射到一个称为 EPC (enclave page cache) 的特殊的物理内存区域. EPC 内的数据是由内存加密引擎 MEE (memory encryption engine) 来进行加密维护的, 进而保证了数据不被运行其飞地的不可信系统、超级管理员以及系统固件所访问. 然而, 这种保护也会受到飞地内存的受限, 例如: SGX1 代预留 128 MB 或 256 MB 加密保护的内存, 现如今第三代英特尔®至强®可扩展处理器开始内置英特尔软件防护扩展, 其安全飞地的容量最多可达单核 CPU 512 GB, 双路共计 1 TB 容量, 可满足更大执行空间需求.

(2) 远程认证: Intel SGX 远程认证允许远程用户 (验证者) 验证飞地 (证明者) 是否被正确创建且安全运行在 SGX 平台上. 在此认证中, 主要涉及注册开发飞地用户 ISV (independent software vendor)、验证飞地的英特尔认证服务验证者 IAS (Intel attestation service) 以及运行 SGX 飞地的平台证明者. 其过程如下.

1) ISV 发送飞地认证质询.

2) 运行 SGX 飞地平台生成包含其飞地度量值 (MRENCLAVE) 的验证报告, 此报告可由 Intel 签名的 QE (quoting enclave) 通过本地认证生成. 同时, QE 通过验证密钥对其签名并发送至 IAS.

3) IAS 验证该签名并利用 Intel 私钥对验证结果进行签名.

4) ISV 通过对比验证结果的签名与飞地的度量值来判断飞地是否可信, 如签名值与飞地度量值不等则验证失败, ISV 将拒绝加载飞地, 反之加载.

(3) 数据封装: 因 SGX 内存受到严格限制, 飞地一旦被创建且实例化, 就会将数据保存在飞地. 当飞地关闭时, 飞地内部的敏感信息就会丢失, 这期间必须保存这些敏感数据以便将来使用. 因此, SGX 提供一种封装技术将其敏感数据封装到飞地之外. 此封装特性是一种只允许检索该飞地平台上生成密钥的机制, 飞地可以使用该密钥对飞地内部敏感数据加密和解密. 如数据需要存储在飞地外时, 则需利用 CPU 硬件派生密钥对数据进行封装.

1.2 数据聚合

数据聚合不仅为用户提供全面且准确的数据统计分析, 还能提高系统分析的效率. 数据拥有者通过加密的方

式来保护其数据的隐私性, 例如医疗数据、交通数据. 因此, 密态数据聚合逐渐成为国内外学者关注的焦点, 各类数据聚合方案层出不穷, 且均要求云服务器不得从聚合结果中还原或获取数据拥有者的敏感信息. Li 等人[18]提出了一种分布式增量数据聚合方案, 通过构建聚合树、聚合路由以最小的开销实现辖区内的数据聚合. Zeng 等人[19]基于 BGN 同态加密提出了可选隐私保护聚合方案, 根据用户数据的敏感度可以选择无隐私加密和隐私加密两种不同加密方式, 从而实现敏感数据的密文聚合和非敏感数据的聚合. Wang 等人[20]提出了基于 fog 的公共云计算匿名安全聚合方案 (anonymous and secure aggregation scheme, ASAS), 通过利用假名来保护终端设备的身份实现匿名聚合. Lu 等人[21]提出了一种面向智能电网通信的高效且隐私保护的聚合方案 (efficient and privacy-preserving aggregation scheme, EPPA). 该方案采用 Paillier 同态加密对结构化数据进行加密, 并实现了密文数据直接在网关进行聚合无需解密, 运营中心即可获得聚合结果. 同时, 还采用批量验证技术降低认证成本. Wang 等人[22]提出了一种基于区块链的边缘计算赋能物联网安全数据聚合策略, 该策略在区块头中集成任务安全级别和安全标签, 并制定新的生成规则来提高系统吞吐量, 又将敏感任务和任务接收者进行分组, 以防止隐私泄露. Li 等人[23]针对物联网场景设计了一种边缘计算辅助数据聚合方案, 既保障了数据机密性及隐私性, 又实现了数据源认证, 并显著降低了通信负担. Fan 等人[24]提出了一种针对内部攻击者的安全智能电网用电数据聚合方案, 具有安全的批量验证功能. Tang 等人[25]提出了一种安全收集多源健康数据的聚合方案, 并采用签名技术保持对患者的激励以及在健康数据中加入噪声, 实现了差分隐私. Han 等人[26]提出了一种基于云辅助的无线体域网的隐私保护和多功能健康数据聚合容错机制, 在不侵犯隐私的情况下通过聚合密文计算用户的多个统计函数. Wang 等人[30]提出了一种基于身份的智能电网数据聚合协议, 不仅可以阻止未经授权的读取, 还可以防止意外错误和抵御恶意篡改攻击. 虽然, 上述方案均采用同态加密来保障数据的机密性, 但在其数据安全管控、细粒度统计分析以及分析效率等方面上仍有不足.

1.3 数据共享

为了实现数据的安全共享, Cheng 等人[31]提出了一种基于区块链的 MCPS 网络模型, 确保了医疗数据无法被篡改, 并建立了完整的追溯机制. Wang 等人[32]基于区块链设计了轻量级数据共享架构, 以保护灾难下的数据共享, 并对行为异常的实体进行不可篡改的跟踪. Mollah 等人[33]提出了一种安全的数据搜索和共享方案, 即在边缘设备上实现弱信任假设, 同时也支持将计算密集型的加密和解密委托给边缘设备. Zhu 等人[34]提出了一种适用于云边协同网络的安全数据共享方案, 实现了 PHC 系统中 IoTS 数据的安全上传、个性密文的归一化、动态多用户的细粒度访问控制、高效的解密操作, 并给出了形式化的安全性证明. 同时引入代理重加密 (proxy re-encryption, PRE) 机制实现了基于单方变换, 即无需对 IoTS 终端进行任何操作的密钥更新. Wang 等人[35]针对区块链赋能的物联网场景, 提出了一种高效可验证的可搜索方案, 实现聚合授权与可信撤销. Xu 等人[36]通过引入一种新的基于属性的可验证外包加解密设计, 提出了一种实用且安全的数据共享系统. Liu 等人[37]提出了一种仅需一个聚合密钥就可以实现文档集上搜索与验证权限共享的搜索方案. Niu 等人[38]基于属性加密、数据去重以及可验证技术, 提出了一种面向智慧医疗场景的云存储数据共享方案. 然而, 在保证数据安全共享的前提下, 从提高计算的效率, 降低用户的计算代价, 实现多种类细粒度统计分析方法的角度看, 这些方案仍有待提高.

2 基础知识

对于一类定义在有限域上的椭圆曲线, 双线性映射能够将曲线上的两个点映射到基域中的一个元素.

定义 1 (双线性映射)[39]. 设 G_1, G_2 和 G_T 是 3 个阶为 p 的循环群, 设 g_1 是群 G_1 的生成元, g_2 是群 G_2 的生成元. 设映射 $e: G_1 \times G_2 \to G_T$ 满足以下 3 个属性.

(1) 双线性: 对于任意两个元素 $\forall u \in G_1$, $v \in G_2$ 和 $a, b \in Z_p$, 使得 $e(u^a, v^b) = e(u, v)^{ab}$.

(2) 非退化性: 存在两个元素 $\forall g_1$, $g_2 \in G_1$ 使得 $e(g_1, g_2) \neq 1 \in G_T$, 其中 1 为 G_T 的单位元.

(3) 可计算性: 对于任何两个 $u \in G_1$ 和 $v \in G_2$, 存在一个多项式时间算法计算 $e(u, v)$, 则称 e 为群 G_1 和 G_2 到群 G_T 上的一个双线性映射. 如果 $G_1 = G_2 = G$ 且 $g_1 = g_2 = g$, 则称 e 为一个对称双线性映射, 否则称 e 为一个非对称

双线性映射.

定义 2 (DBDH 假设问题). 设 G 和 G_T 是均为 p 阶的循环群, g 为群 G 的生成元. 设映射 $e: G\times G \to G_T$. 对于任意概率多项式时间内的敌手 $\mathcal{A}$, 给定 (g, p, e, g^a, g^b, g^c), 敌手的优势 $|\Pr[\mathcal{A}(g,p,e,g^a,g^b,g^c,h_t)=b]-1/2|$ 可以忽略. 其中, $h_0 = e(g,g)^{abc}$、$h_1 = e(g,g)^{\varpi}$、$a,b,c,\varpi \in Z_p$ 和 $t\in\{0,1\}$.

定义 3 (计算性 Diffie-Hellman 问题). 设 G 是一个阶为大素数 p 的循环群, 设 g 为循环群 G 的生成元. 给定 $(g, g^{\alpha}, g^{\beta}) \in G^3$, 其中 $\alpha,\beta \in Z_p$, 计算 $g^{\alpha\beta} \in G$.

定义 4 (计算性 Diffie-Hellman 假设). 设任何概率多项式时间算法 $\mathcal{T}$ 成功求解群 G 上的计算性 Diffie-Hellman 问题的优势为:

$$\mathcal{ADV}_{\mathrm{CDH}}(\mathcal{T}) = \Pr\left[\mathcal{T}(q,g,g^{\alpha},g^{\beta}) = g^{\alpha\beta} | g\in G, \alpha,\beta\in Z_p\right].$$

计算性 Diffie-Hellman 假设定义为: 对于任何概率多项式时间算法 $\mathcal{T}$, $\mathcal{ADV}_{\mathrm{CDH}}(\mathcal{T})$ 是可忽略的.

2.1 哈希函数

哈希函数是将任意长度的输入映射到一个固定长度的输出上[40]. 一般表示形式为 $H:\{0,1\}^* \to \{0,1\}^{\alpha}$, 其中 α 代表输出内容的长度. 通常来讲, 哈希函数具有下列性质[41].

(1) 确定性: 对于同一输入 m, 哈希函数的输出 $H(m)$ 总是相同的.

(2) 单向性: 对于一个哈希输出结果 $H(m)$, 多项式内无法有效逆向计算出 m 值.

(3) 抗碰撞性: 对于一个哈希函数的结果 $H(m)$, 不存在任何有效算法可以在多项式时间内找到一个 n, 满足 $H(n)=H(m)$ 且 $n\neq m$.

2.2 新型 BGN 同态加密算法

BGN 同态加密算法[17] 属于全同态加密技术中的一种, 其特征在于可以进行多次加法同态和一次乘法同态加密, 文献 [42] 将 BGN[17] 密码体制扩展到双消息加密的新型 BGN 同态加密算法, 具体的细节如下.

Step 1. $KeyGen(1^{\lambda}) \to (pk, sk)$. 此算法输入安全参数 λ, 输出为公钥 pk 和私钥 $sk=\{Sk_1, Sk_2\}$. 首先, 选择 4 个长度均为 λ 的大素数 p_0'、p_1'、p_2' 和 p_3', 设置 $N = p_0'p_1'p_2'p_3'$. 选取 $\tilde{G}_1$ 和 $\tilde{G}_2$ 两个阶为 N 循环群, g_1 是 $\tilde{G}_1$ 循环群的生成元. $e:\tilde{G}_1\times\tilde{G}_1 \to \tilde{G}_2$ 为循环群 $\tilde{G}_1$ 和 $\tilde{G}_2$ 的双线性映射. 设置 $u_1' = g_1^{p_1'p_2'}$、$u_2' = g_1^{p_0'p_2'}$ 和 $u_3' = g_1^{p_0'p_1'}$. 公钥 $pk = \{N, \tilde{G}_1, \tilde{G}_2, e, g_1, u_1', u_2', u_3'\}$, 私钥 $Sk_1 = p_1'p_2'p_3'$ 和 $Sk_2 = p_0'p_2'p_3'$. 其中, u_1' 是 $\tilde{G}_1$ 的 $p_0'p_3'$ 阶循环子群的生成元, u_2' 是 $\tilde{G}_1$ 的 $p_1'p_3'$ 阶循环子群的生成元, u_3' 是 $\tilde{G}_1$ 的 $p_2'p_3'$ 阶循环子群的生成元.

Step 2. $Enc(m, m^2, pk, r) \to C$. 此算法输入明文消息空间 $m\in[1,T_1]$ 和 $m^2\in[1,T_2]$, 其中 $T_1 < p_0'p_3'$, $T_2 < p_1'p_3'$, 公钥 pk 以及随机数 $r\in Z_N$, 输出则为密文 $C = (u_1')^m (u_2')^{m^2} (u_3')^r \in \tilde{G}_1$.

Step 3. $Dec(C, sk) \to \{m, m^2\}$. 此算法输入为密文 $C = (u_1')^m (u_2')^{m^2} (u_3')^r \in \tilde{G}_1$、私钥 $Sk_1 = p_1'p_2'p_3'$ 和 $Sk_2 = p_0'p_2'p_3'$, 输出为明文 $\{m, m^2\}$. 具体解密过程为 $C^{Sk_1} = ((u_1')^m (u_2')^{m^2} (u_3')^r)^{p_1'p_2'p_3'} = ((u_1')^{p_1'p_2'p_3'})^m$, 接下来利用文献 [17] 中提及的 Pollard 的 lambda 解密方法, 通过计算以 $g_1^{(p_1')^2(p_2')^2p_3}$ 为底 C^{Sk_1} 的离散对数便可获得明文消息 m. 同理, 再计算 $C^{Sk_2} = ((u_1')^m (u_2')^{m^2} (u_3')^r)^{p_0'p_2'p_3'} = ((u_2')^{p_0'p_2'p_3'})^{m^2}$ 便可获得明文消息 m^2.

3 基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案

为了实现医疗密文数据的安全聚合与统计分析, 本节首先设计并提出 ODMT-BGN 算法, 然后详细介绍基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案的系统模型、威胁模型和方案详细描述.

3.1 ODMT-BGN 算法

新型 BGN 同态加密算法存在消息空间利用率过低的问题, 因此, 本节引入膨胀因子概念, 膨胀因子为密文长度与明文长度的比值.

由第 2.2 节 Step 1 中 $KeyGen$ 算法, 可知群 $\tilde{G}_1$ 的阶 $N = p_0'p_1'p_2'p_3'$, 其中要求消息 $m < p_0'p_3'$ 且 $m^2 < p_1'p_3'$, 通过公式 (1) 可推导出 m 必定满足 $m \leqslant \left\lfloor \sqrt{p_1'p_3'} \right\rfloor$. 由第 2.2 节的 Step 2 可知密文 $C = (u_1')^m (u_2')^{m^2} (u_3')^r \in \tilde{G}_1$, 故而得到一个关

系映射 Map, 如公式 (2) 所示:

$$\begin{cases} m < p'_0p'_3 \\ m^2 < p'_1p'_3 \end{cases} \Rightarrow \begin{cases} m < p'_0p'_3 \\ m \leqslant \left\lfloor \sqrt{p'_1p'_3} \right\rfloor \end{cases} \tag{1}$$

$$Map \begin{cases} Z_{\left\lfloor \sqrt{p'_1p'_3} \right\rfloor} \times Z_N \mapsto \tilde{G}_1 \\ (m, m^2, r) \mapsto (u'_1)^m (u'_2)^{m^2} (u'_3)^r \bmod N \end{cases} \tag{2}$$

又因为 p'_0、p'_1、p'_2 和 p'_3 是长度均为 λ 的大素数, 因此, 可计算出膨胀因子 Cef, 见公式 (3), 可看出群 $\tilde{G}_1$ 中有一部分空间没有得到充分利用. 下面给出详细的解释, 其中 $\left|\tilde{G}_1\right|$ 表示群的阶数, $|p'|$ 表示取长度.

$$Cef = \log_2 \frac{\left|\tilde{G}_1\right|}{\left|Z_{\left\lfloor \sqrt{p'_1p'_3} \right\rfloor}\right|} = \frac{\left|p'_0p'_1p'_2p'_3\right|}{\left\|\left\lfloor \sqrt{p'_1p'_3} \right\rfloor\right\|} \approx 4/1 \tag{3}$$

设 $\left\lfloor \sqrt{p'_1p'_3} \right\rfloor \leqslant a < p'_0p'_3, 0 \leqslant b < p'_1p'_3, c \in Z_N$, 则 $\forall m \in [0, \left\lfloor \sqrt{p'_1p'_3} \right\rfloor]$, $r \in Z_N$, 有 $Enc(m, m^2, r) \neq (u'_1)^a (u'_2)^b (u'_3)^c \bmod N \in \tilde{G}_1$.

因此, 为了使用更充足的消息空间, ODMT-BGN 算法设 $|N| \approx |N'|$ 且 $N' = p_0p_1p_2p_3$, 其中 p_0、p_1、p_2 和 p_3 都是大素数, 但需满足 $|p_0| : |p_1| : |p_2| : |p_3| = 1 : 3 : 1 : 1$ 的条件, 则计算密文膨胀因子 Cef', 如公式 (4) 所示:

$$Cef' = \frac{|p_0p_1p_2p_3|}{\left\|\left\lfloor \sqrt{p_1p_3} \right\rfloor\right\|} \approx 6/2 = 3/1 \tag{4}$$

由此可得, 在针对 m 和 m^2 的加密情况下, ODMT-BGN 算法降低了密文膨胀因子, 即扩展了消息空间, 使得 ODMT-BGN 算法更为安全. ODMT-BGN 算法详细描述如下.

(1) $KeyGen(1^\lambda) \to (pk, s)$. 此算法输入安全参数 λ, 输出为公钥 mpk_1 和私钥 $s = \{s_1, s_2\}$. 首先, 选择 4 个大素数 p_0、p_1、p_2 和 p_3, 并设置 $N' = p_0p_1p_2p_3$ 且满足 $|p_0| : |p_1| : |p_2| : |p_3| = 1 : 3 : 1 : 1$, $p_i > 2^{\lambda/2}, i \in \{0,1,2,3\}$. 其中, 要求 $|N'| \approx |N|$ (其中 N 为合数阶, 见第 2.2 节), G_1 和 G_2 为 N' 阶的循环群, g_1 是 G_1 循环群的生成元, $e : G_1 \times G_1 \to G_2$ 为循环群 G_1 和 G_2 的双线性映射, 同时计算 $u_1 = g_1^{p_1p_2}$、$u_2 = g_1^{p_0p_2}$ 和 $u_3 = g_1^{p_0p_1}$. 输出公钥为 $mpk_1 = \{N', G_1, G_2, e, g_1, u_1, u_2, u_3\}$, 私钥为 $s_1 = p_1p_2p_3$ 和 $s_2 = p_0p_2p_3$. 其中, u_1 是 G_1 的 p_0p_3 阶循环子群的生成元, u_2 是 G_1 的 p_1p_3 阶循环子群的生成元, u_3 是 G_1 的 p_2p_3 阶循环子群的生成元.

(2) $Enc(m, m^2, mpk_1, r) \to C$. 此算法输入明文消息空间 $m \in [1, \tau_1]$ 和 $m^2 \in [1, \tau_2]$、公钥 mpk_1 以及随机数 $r \in Z_{N'}$, 其中 $\tau_1 < p_0p_3$, $\tau_2 < p_1p_3$. 输出则为密文 $C = u_1^m u_2^{m^2} u_3^r \in G_1$.

(3) $Dec(C, s) \to \{m, m^2\}$. 此算法输入为密文 $C = u_1^m u_2^{m^2} u_3^r$、私钥 $s_1 = p_1p_2p_3$ 和 $s_2 = p_0p_2p_3$, 输出为明文 $\{m, m^2\}$. 解密过程为 $C^{s_1} = (u_1^m u_2^{m^2} u_3^r)^{p_1p_2p_3} = (u_1^{p_1p_2p_3})^m$, 随后, 利用文献 [17] 中提及的 Pollard 的 lambda 解密方法, 计算以 $g_1^{(p'_1)^2(p'_2)^2p_3}$ 为底 C^{Sk_1} 的离散对数便可得到明文消息 m. 同理, 计算 $C^{Sk_2} = ((u'_1)^m (u'_2)^{m^2} (u'_3)^r)^{p'_0p'_2p'_3} = ((u'_2)^{p'_0p'_2p'_3})^{m^2}$ 便可得到明文消息 m^2.

3.2 系统模型

本方案主要包含 5 类实体, 分别是密钥分发中心 $\mathcal{KDC}$、数据拥有者 $\mathcal{DO}$、边缘服务器 $\mathcal{ES}$、云服务器 $\mathcal{CS}$ 和研究中心 $\mathcal{RC}$. $\mathcal{RC}$ 与 $\mathcal{CS}$ 具有可信硬件 TEE. 如图 2 所示, 方案中各个实体任务与系统流程描述如下.

(1) 密钥分发中心 $\mathcal{KDC}$: 负责为系统中实体生成密钥以及公共参数, 随后进行密钥分发以及公开系统公共参数.

(2) 数据拥有者 $\mathcal{DO}$: 主要利用 $\mathcal{KDC}$ 分发的密钥对医疗数据进行加密, 再对密文进行签名, 并将密文与签名一起上传至所属的边缘服务器.

(3) 边缘服务器 $\mathcal{ES}$: 主要负责聚合验证辖区内所有 $\mathcal{DO}$ 的签名及聚合 $\mathcal{DO}$ 上传的密文, 并对聚合后的密文进行签名, 然后将聚合密文与聚合签名上传至云服务器. 此外, $\mathcal{ES}$ 还需要对 $\mathcal{RC}$ 的访问请求进行授权操作.

(4) 云服务器 $\mathcal{CS}$: 分为存储和 TEE 两部分. 其中 TEE 主要负责对所有 $\mathcal{ES}$ 上传签名进行聚合验证以及根据分析请求进行统计分析操作. 此外, TEE 也负责对分析结果进行加密与签名, 再将分析结果和签名发送给 $\mathcal{RC}$. $\mathcal{CS}$ 的

存储部分则是存储所有 $\mathcal{ES}$ 上传的密文数据.

(5) 研究中心 $\mathcal{RC}$: 首先与 $\mathcal{CS}$ 进行远程认证来确保 TEE 安全运行在云服务器 $\mathcal{CS}$ 上, 其次负责向 $\mathcal{ES}$ 发送授权请求并聚合授权密钥. 最后, 还负责发送分析请求给 $\mathcal{CS}$, 并对 $\mathcal{CS}$ 发送分析结果进行验证与解密.

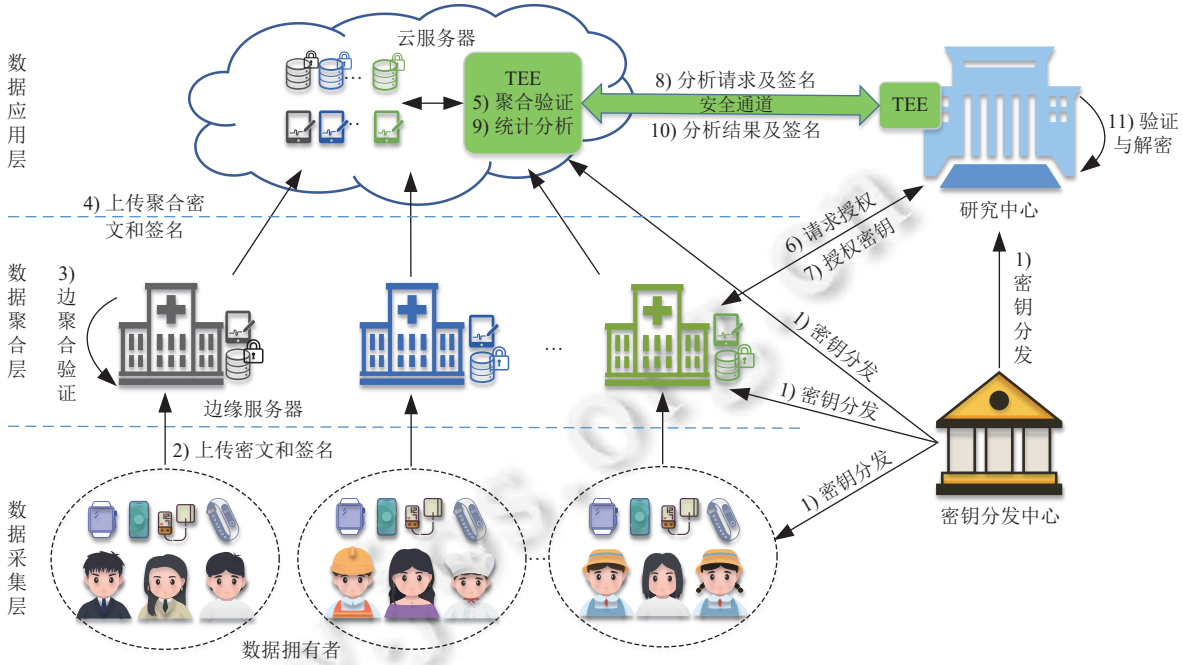


图 2 系统模型

如图 2 所示, 系统流程如下.

1) 密钥分发中心 $\mathcal{KDC}$ 为系统中数据拥有者、边缘服务器、云服务器以及研究中心等实体进行密钥分发;

2) 数据拥有者 $\mathcal{DO}$ 通过系统生成的公钥对医疗数据进行加密, 并对密文进行数字签名, 然后再将其密文和签名上传至对应的边缘服务器 $\mathcal{ES}$;

3) 边缘服务器 $\mathcal{ES}$ 接收辖区内所有 $\mathcal{DO}$ 的密文和签名, 先对所有签名进行边级的聚合验证. 若验证通过, 则对所有密文进行聚合得到边级聚合密文, 并对边级聚合密文进行数字签名, 否则, 将逐个验证 $\mathcal{DO}$ 的签名以确定哪些 $\mathcal{DO}$ 的医疗数据被篡改;

4) 所有边缘服务器 $\mathcal{ES}$ 将边级聚合密文和边级签名上传至云服务器 $\mathcal{CS}$;

5) 云服务器 $\mathcal{CS}$ 先对所有 $\mathcal{ES}$ 上传的边级签名进行聚合验证, 若验证通过, 则对所有边级聚合密文进行聚合得到云级聚合密文, 否则, 将逐个验证 $\mathcal{ES}$ 的边级签名以确定哪些 $\mathcal{ES}$ 的医疗数据被篡改;

6) 研究中心 $\mathcal{RC}$ 向边缘服务器 $\mathcal{ES}$ 发送医疗数据访问授权请求;

7) $\mathcal{ES}$ 响应研究中心 $\mathcal{RC}$ 访问授权操作, 并生成授权密钥, 再将加密的授权密钥及其签名返还给 $\mathcal{RC}$;

8) $\mathcal{RC}$ 接收加密的授权密钥及其签名, 然后对其签名进行验证, 若验证通过, 则对加密的授权密钥进行解密并聚合授权密钥, 再利用聚合密钥生成分析请求并对其密文进行数字签名; 否则重新发起授权请求; 最后再将分析请求及签名发送至 $\mathcal{CS}$;

9) $\mathcal{CS}$ 接收分析请求及其签名, 并对其签名进行验证, 如验证通过, 则在 TEE 内根据解密后的分析请求执行统计分析并得到相应的分析结果;

10) $\mathcal{CS}$ 将分析结果进行加密和签名, 并返还给 $\mathcal{RC}$;

11) $\mathcal{RC}$ 接收密文分析结果和签名, 并对签名进行验证, 若验证通过, 再对密文分析结果进行解密得到最终的

分析结果.

3.3 威胁模型

在系统模型中, 我们假设云服务器 $\mathcal{CS}$ 的 TEE、密钥分发中心 $\mathcal{KDC}$、数据拥有者 $\mathcal{DO}$ 以及研究中心 $\mathcal{RC}$ 是可信的, 边缘服务器 $\mathcal{ES}$ 是诚实且好奇的, 即 $\mathcal{ES}$ 会诚实地执行协议, 却可能试图获取密文中包含的敏感信息. 而 $\mathcal{CS}$ 存储部分和外部敌手是不可信的. 换句话说, 敌手通过观察 $\mathcal{DO}$ 与 $\mathcal{ES}$、$\mathcal{ES}$ 与 $\mathcal{CS}$、$\mathcal{RC}$ 与 $\mathcal{CS}$ 和 $\mathcal{RC}$ 与 $\mathcal{ES}$ 之间的通信获得传输信息, 并企图通过篡改密文信息破坏数据完整性来获取明文信息. 此外, 边缘服务器之间也可能会对彼此上传的数据好奇. 因此, 还需假设所有实体之间不会共谋. 本节定义安全性需求如下.

(1) 正确性: 方案中每个算法都正确地执行, 密文的分析结果将永远能被研究中心正确验证与解密;

(2) 机密性: 除数据拥有者本身和 TEE 之外, 在不知道密钥的前提下, 边缘服务器、云服务器存储部分、外部敌手等实体都无法根据密文数据得到有关明文数据的任何消息, 且云服务器和边缘服务器只能得到对应的密文聚合结果;

(3) 完整性: 云服务器的 TEE 和边缘服务器应该检测待聚合的密文数据是否被篡改.

3.4 方案详细描述

(1) 初始化阶段

假设系统存在 n 个边缘服务器 $\mathcal{ES}_i$, 每个边缘服务器下存在 β 个数据拥有者 $\mathcal{DO}_{i_j}$, $j=1,\ldots,\beta$, 一个云服务器 $\mathcal{CS}$、一个研究中心 $\mathcal{RC}$ 和一个密钥分发中心 $\mathcal{KDC}$. $\mathcal{KDC}$ 为每个实体生成为唯一的身份标识. 其中, $id_{\mathcal{ES}_i}$ 代表第 i 个边缘服务器的身份标识, $id_{\mathcal{DO}_{i_j}}$ 代表第 i 个 $\mathcal{ES}_i$ 下第 j 个数据拥有者的身份标识, $id_{\mathcal{CS}}$ 代表云服务器的身份标识, $id_{\mathcal{RC}}$ 代表研究中心的身份标识. 详细步骤如下.

1) $\mathcal{KDC}$ 根据输入的安全参数 λ, 采用 ODMT-BGN 算法中 *KeyGen* 算法生成的公钥 mpk_1 和私钥 (s_1,s_2), 其中 $mpk_1=(G_1,G_2,g_1,u_1,u_2,u_3,e,N')$, $s_1=p_1p_2p_3$ 和 $s_2=p_0p_2p_3$.

2) $\mathcal{KDC}$ 再根据安全参数 λ 生成另一个双线性映射 $\tilde{e}:\mathcal{G}_1\times\mathcal{G}_1\to\mathcal{G}_2$, 其中, 循环群 $\mathcal{G}_1$, $\mathcal{G}_2$ 的阶均为大素数 q 且 $q>2^{\lambda}$, g 是群 $\mathcal{G}_1$ 的生成元, g_t 是群 $\mathcal{G}_2$ 的生成元. 然后, $\mathcal{KDC}$ 选取一个随机数 $\delta\leftarrow Z_q^*$ 作为主密钥, 并计算 $y=g^{\delta}$, $\mathcal{W}=\tilde{e}(H_1(id_{\mathcal{RC}}),y)$, $\mathcal{W}_i=\tilde{e}(H_1(id_{\mathcal{ES}_i}),y)$. 此外, $\mathcal{KDC}$ 设置两个抗碰撞哈希函数 $H_1:\{0,1\}^*\to\mathcal{G}_1$ 和 $H_2:\{0,1\}^*\to\mathcal{G}_1$. 随后, $\mathcal{KDC}$ 再分别为 $\mathcal{CS}$、$\mathcal{RC}$、$\mathcal{ES}_i$ 及 $\mathcal{DO}_{i_j}$ 生成私钥 $sk_{\mathcal{CS}}=H_1(id_{\mathcal{CS}})^{\delta}$, $sk_{\mathcal{RC}}=H_1(id_{\mathcal{RC}})^{\delta}$, $sk_i=H_1(id_{\mathcal{ES}_i})^{\delta}$ 和 $sk_{i_j}=H_1(id_{\mathcal{DO}_{i_j}})^{\delta}$.

3) $\mathcal{KDC}$ 根据安全参数 λ 再生成一个乘法循环群 $\mathcal{G}_3$ 的阶为大素数 q, 且 $q>2^{\lambda}$, g_3 是群 $\mathcal{G}_3$ 生成元, $\mathcal{KDC}$ 随机选取随机数 $\alpha_i\leftarrow Z_q^*, 1\leqslant i\leqslant n$. 同时, $\mathcal{KDC}$ 为每个 $\mathcal{ES}_i$ 选取一个对应的聚合密文索引标签 $i\in\{1,\ldots,n\}$ 并设置授权标签集合为 $S=\{1,\ldots,n\}$. 此外, $\mathcal{KDC}$ 设置一个抗碰撞哈希函数 $H:\{0,1\}^*\to\mathcal{G}_3$.

4) 最后, $\mathcal{KDC}$ 发布系统公共参数 $mpk_1=(G_1,G_2,g_1,u_1,u_2,u_3,e,N')$, $mpk_2=(\mathcal{G}_1,\mathcal{G}_2,g,g_t,H_1,H_2,\tilde{e},y,q,\mathcal{W}_i,\mathcal{W})$ 和 $mpk_3=(\mathcal{G}_3,g_3,H,q,S)$. $\mathcal{KDC}$ 将 $\{sk_{\mathcal{RC}},s_1,s_2,id_{\mathcal{RC}}\}$, $\{sk_i,\alpha_i,id_{\mathcal{ES}_i}\}$, $\{sk_{\mathcal{CS}},id_{\mathcal{CS}}\}$ 和 $\{sk_{i_j},id_{\mathcal{DO}_{i_j}}\}$ 分别发给对应的 $\mathcal{RC}$、$\mathcal{ES}_i$、$\mathcal{CS}$ 和 $\mathcal{DO}_{i_j}$, 其中 $1\leqslant i\leqslant n$, $1\leqslant j\leqslant\beta$. 特别地, $\mathcal{RC}$ 通过安全信道将 $\{s_1,s_2\}$ 共享到 $\mathcal{CS}$ 的 TEE 中.

(2) 密文数据上传阶段

1) $\mathcal{DO}_{i_j}$ 选取一个随机数 $r_{i_j}\leftarrow Z_{N'}$, 并用 ODMT-BGN 算法对 m_{i_j} 进行加密, 得到密文 $c_{i_j}=u_1^{m_{i_j}}u_2^{m_{i_j}^2}u_3^{r_{i_j}}\in G_1$.

2) $\mathcal{DO}_{i_j}$ 选取一个随机数 $w_{i_j}\leftarrow Z_q^*$, 利用签名私钥 sk_{i_j} 对密文 c_{i_j} 进行签名, 得到签名 $\sigma_{i_j}=(\sigma_{i_j,1},\sigma_{i_j,2})=(g^{w_{i_j}}, sk_{i_j}\cdot H_2(c_{i_j}\|t_{i_j})^{w_{i_j}})$, 其中 t_{i_j} 代表每个 $\mathcal{DO}_{i_j}$ 的时间戳.

3) 每个 $\mathcal{DO}_{i_j}$ 将这些可验证的数据信息 $\psi_{i_j}=\{c_{i_j},\sigma_{i_j},t_{i_j},id_{\mathcal{DO}_{i_j}}\}$, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant\beta$ 发送给对应的 $\mathcal{ES}_i$.

(3) 边聚合验证与密文数据上传阶段

1) 每个 $\mathcal{ES}_i$ 接收辖区内所有 $\mathcal{DO}_{i_j}$ 发送的 ψ_{i_j}, 并对所有 ψ_{i_j} 执行公式 (5) 的批量验证, 其中 $\varsigma_{i_j}\leftarrow Z_q^*$.

$$\tilde{e}\left(g,\prod_{j=1}^{\beta}(\sigma_{i_j,2})^{\varsigma_{i_j}}\right)=\tilde{e}\left(y,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}}))^{\varsigma_{i_j}}\right)\prod_{j=1}^{\beta}\tilde{e}\left(\sigma_{i_j,1},(H_2(c_{i_j}\|t_{i_j}))^{\varsigma_{i_j}}\right)\tag{5}$$

如验证未通过, 则表明至少有一个 $\mathcal{DO}_{i_j}$ 的 ψ_{i_j} 在传输过程中被替换或篡改. 因此, $\mathcal{ES}_i$ 需要对每个 $\mathcal{DO}_{i_j}$ 的 ψ_{i_j}

进行公式 (6) 的逐个验证, 直至所有验证通过并对所有密文进行存储.

$$\tilde{e}\left(g,\sigma_{i_j,2}\right)=\tilde{e}\left(y,H_1(id_{\mathcal{DO}_{i_j}})\right)\tilde{e}\left(\sigma_{i_j,1},H_2(c_{i_j}\|t_{i_j})\right) \tag{6}$$

2) 验证通过后, 每个 $\mathcal{ES}_i$ 对辖区内所有密文 c_{i_j} 进行聚合得 $c_i=\prod_{j=1}^{\beta}c_{i_j}=u_1^{\sum_{j=1}^{\beta}m_{i_j}}u_2^{\sum_{j=1}^{\beta}m_{i_j}^2}u_3^{\sum_{j=1}^{\beta}r_{i_j}}$. 同时, 每个 $\mathcal{ES}_i$ 选取随机数 $w_i,\vartheta_i\leftarrow Z_q^*$ 并计算 $V_i=g_3^{\vartheta_i}$, 以及计算 $\mathcal{ES}_i$ 所属的密文数据标签 $ct_i=\vartheta_i+\alpha_i\cdot H(i)$, $1\leqslant i\leqslant n$. 随后, $\mathcal{ES}_i$ 分别对 c_i 与 ct_i 进行数字签名得到 $\sigma_i=(\sigma_{i,1},\sigma_{i,2})=(g^{w_i},sk_i\cdot H_2(c_i\|t_i)^{w_i})$ 和 $\tilde{\sigma}_i=(\tilde{\sigma}_{i,1},\tilde{\sigma}_{i,2})=(g^{w_i},sk_i\cdot H_2(ct_i\|c_i\|V_i\|t_i)^{w_i})$, t_i 代表每个 $\mathcal{ES}_i$ 的时间戳.

3) 每个 $\mathcal{ES}_i$ 发送可验证数据信息 $\psi_i=\{c_i,\sigma_i,ct_i,\tilde{\sigma}_i,V_i,id_{\mathcal{ES}_i},t_i\}$, $1\leqslant i\leqslant n$ 给云服务器 $\mathcal{CS}$.

(4) 云聚合验证阶段

$\mathcal{CS}$ 接收每个 $\mathcal{ES}_i$ 发送的可验证信息 ψ_i 进行公式 (7) 批量验证, 其中 $\varsigma_i\leftarrow Z_q^*$. 如验证未通过, 则需要逐个执行公式 (8) 验证来判断哪个 $\mathcal{ES}_i$ 上传的数据被替换或篡改. 同理, 如果公式 (9) 验证未通过, 则 $\mathcal{CS}$ 执行公式 (10) 验证来判断哪个 $\mathcal{ES}_i$ 上传密文数据的标签被替换或篡改.

$$\tilde{e}\left(g,\prod_{i=1}^{n}(\sigma_{i,2})^{\varsigma_i}\right)=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\sigma_{i,1},(H_2(c_i\|t_i))^{\varsigma_i}) \tag{7}$$

$$\tilde{e}(g,\sigma_{i,2})=\tilde{e}(y,H_1(id_{\mathcal{ES}_i}))\tilde{e}(\sigma_{i,1},H_2(c_i\|t_i)) \tag{8}$$

$$\tilde{e}\left(g,\prod_{i=1}^{n}(\tilde{\sigma}_{i,2})^{\varsigma_i}\right)=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\tilde{\sigma}_{i,1},(H_2(ct_i\|c_i\|V_i\|t_i))^{\varsigma_i}) \tag{9}$$

$$\tilde{e}(g,\tilde{\sigma}_{i,2})=\tilde{e}(y,H_1(id_{\mathcal{ES}_i}))\tilde{e}(\tilde{\sigma}_{i,1},H_2(ct_i\|c_i\|V_i\|t_i)) \tag{10}$$

(5) 授权阶段

1) $\mathcal{RC}$ 生成请求信息集合 $\mathcal{I}\subseteq S$, 然后, $\mathcal{RC}$ 随机选取 $\gamma_i,\upsilon_i\leftarrow Z_q^*$ 对集合 $\mathcal{I}$ 中任意一个 $\mathcal{I}_i$ 进行加密和签名, 分别得 $\mathcal{CI}_i$ 和 $\hat{\sigma}_i$, 再将可验证信息 $\hat{\psi}_i=\{\mathcal{CI}_i,\hat{\sigma}_i,\hat{t}_i,id_{\mathcal{RC}}\}$ 发给对应的 $\mathcal{ES}_i$, 其中 $\hat{t}_i$ 代表 $\mathcal{RC}$ 的时间戳.

$$\mathcal{CI}_i=(\mathcal{CI}_{i,1},\mathcal{CI}_{i,2})=\left(g^{\gamma_i},g_t^{\mathcal{I}_i}\cdot\mathcal{W}_i^{\gamma_i}\right)$$

$$\hat{\sigma}_i=(\hat{\sigma}_{i,1},\hat{\sigma}_{i,2})=\left(g^{\upsilon_i},sk_{\mathcal{RC}}\cdot H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}\|\hat{t}_i)^{\upsilon_i}\right)$$

2) $\mathcal{ES}_i$ 对 $\hat{\sigma}_i$ 进行公式 (11) 的验证. 若验证通过, 则对 $\mathcal{CI}_i$ 进行解密得到 $\mathcal{I}_i=\mathcal{CI}_{i,2}/\tilde{e}(sk_i,\mathcal{CI}_{i,1})$. $\mathcal{ES}_i$ 再通过 $\mathcal{I}_i\triangleq i$ 等式判断 $\mathcal{RC}$ 的请求信息 $\mathcal{I}_i$ 是否与 $\mathcal{ES}_i$ 对应的密文索引标签 i 相等. 若相等, $\mathcal{ES}_i$ 计算授权密钥 $x_{\mathcal{I}_i}=g_3^{\alpha_{\mathcal{I}_i}}$, 再以同样的方式发送授权密钥给 $\mathcal{RC}$. 当 $\mathcal{RC}$ 接收所有数据后进行验证和解密得到所有授权密钥 $x_{\mathcal{I}_i}$, 再计算聚合授权密钥 $\mathcal{AK}=\prod_{\mathcal{I}\subseteq S}x_i=\prod_{\mathcal{I}\subseteq S}g_3^{\alpha_i}$.

$$\tilde{e}(g,\hat{\sigma}_{i,2})=\tilde{e}(y,H_1(id_{\mathcal{RC}}))\tilde{e}(\hat{\sigma}_{i,1},H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}\|\hat{t}_i)) \tag{11}$$

(6) 统计分析阶段

1) $\mathcal{RC}$ 利用 $\mathcal{AK}$ 生成加密分析请求 $\mathcal{C}_{\mathcal{Q}}=\mathcal{AK}^{H(\mathcal{Q})}$ 和签名 $\sigma_{\mathcal{Q}}=(\sigma_{\mathcal{Q},1},\sigma_{\mathcal{Q},2})=(g^{\upsilon},sk_{\mathcal{RC}}\cdot H_2(\mathcal{C}_{\mathcal{Q}}\|t_{\mathcal{Q}}\|\mathcal{AM})^{\upsilon})$, $\mathcal{Q}\subseteq S$, $\upsilon\leftarrow Z_q^*$, 其 $t_{\mathcal{Q}}$ 代表 $\mathcal{RC}$ 的时间戳, $\mathcal{AM}=\{Ave,Sum,Var,SD,Cov\}$ 代表统计分析方法, 并将可验证信息 $\psi_{\mathcal{Q}}=\{\mathcal{C}_{\mathcal{Q}},\sigma_{\mathcal{Q}},t_{\mathcal{Q}},id_{\mathcal{RC}},\mathcal{AM}\}$ 发送给 $\mathcal{CS}$ 的 TEE.

2) $\mathcal{CS}$ 中的 TEE 通过公式 (12) 来验证分析请求是否被篡改或替换. 如果验证通过, TEE 则再通过验证 $\prod_{\mathcal{Q}\subseteq S}V_{\mathcal{Q}}\cdot\mathcal{C}_{\mathcal{Q}}\triangleq\prod_{\mathcal{Q}\subseteq S}g_3^{ct_{\mathcal{Q}}}$ 得到分析请求 $\mathcal{Q}$ 对应的密文 $c_{\mathcal{Q}}$.

$$\tilde{e}(g,\sigma_{\mathcal{Q},2})=\tilde{e}(y,H_1(id_{\mathcal{RC}}))\tilde{e}(\sigma_{\mathcal{Q},1},H_2(\mathcal{C}_{\mathcal{Q}}\|t_{\mathcal{Q}}\|\mathcal{AM})) \tag{12}$$

3) TEE 对 $c_{\mathcal{Q}}$ 进行聚合得聚合密文 $C=\prod_{\mathcal{Q}\subseteq S}c_{\mathcal{Q}}=u_1^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}}u_2^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}^2}u_3^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}r_{\mathcal{Q}_j}}$, $\mathcal{Q}\subseteq S$, $j\in[1,\beta]$. 再利用 s_1 和 s_2 进行解密, 分别得到统计和 $R_1=C^{s_1}=\left(u_1^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}}u_2^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}^2}u_3^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}r_{\mathcal{Q}_j}}\right)^{p_1p_2p_3}=g_1^{(p_1^2p_2^2p_3)\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}}$ 以及医疗数据的平方和 $R_2=C^{s_2}=\left(u_1^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}}u_2^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}^2}u_3^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}r_{\mathcal{Q}_j}}\right)^{p_0p_2p_3}=g_1^{(p_0^2p_2^2p_3)\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}^2}$. TEE 再利用 R_1 和 R_2 计算方差、期望或标准差等统计结果 $\mathcal{R}$. 然后, TEE 利用密钥对 $\mathcal{R}$ 进行加密与签名, 得到密文统计结果 $C_{\mathcal{R}}=(c_{\mathcal{R},1},c_{\mathcal{R},2})=(g^{r'},g_t^{\mathcal{R}}\cdot\tilde{e}(H_1(id_{\mathcal{RC}}),y^{r'}))$ 与签名 $\sigma_{\mathcal{R}}=(\sigma_{\mathcal{R},1},\sigma_{\mathcal{R},2})=\left(g^{\varepsilon},sk_{\mathcal{CS}}\cdot H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})^{\varepsilon}\right)$, 其中 $r',\varepsilon\leftarrow Z_q^*$. 利用文献 [17] 中提

及的 Pollard 的 lambda 解密方法, 计算以 $g_1^{p_1^2p_2^2p_3}$ 底的 R_1 的离散对数便可得到统计和 $\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}$, 同理计算以 $g_1^{p_0^2p_2^2p_3}$ 底的 R_2 的离散对数便可得到平方和 $\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta}m_{\mathcal{Q}_j}^2$. 最后, $\mathcal{CS}$ 返回给 $\mathcal{RC}$ 可验证的数据信息 $\psi=\{C_{\mathcal{R}},\sigma_{\mathcal{R}},t_{\mathcal{R}},id_{\mathcal{CS}}\}$.

(7) 验证与解密阶段

1) $\mathcal{RC}$ 接收可验证数据信息 ψ, 并执行公式 (13) 来验证统计分析结果在传输过程中是否被篡改或替换.

$$\tilde{e}(g,\sigma_{\mathcal{R},2})=\tilde{e}(y,H_1(id_{\mathcal{CS}}))\tilde{e}(\sigma_{\mathcal{R},1},H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})) \tag{13}$$

2) 若验证通过, 则 $\mathcal{RC}$ 利用 $sk_{\mathcal{RC}}$ 对密文分析结果 $C_{\mathcal{R}}=(c_{\mathcal{R},1},c_{\mathcal{R},2})=(g^{r'},g_t^R\cdot\mathcal{W}^{r'})$ 进行解密, 并设置 $R=\dfrac{c_{\mathcal{R},2}}{\tilde{e}(sk_{\mathcal{RC}},c_{\mathcal{R},1})}=g_t^{\mathcal{R}}$, 再计算以 g_t 为底 R 的离散对数得到统计结果 $\mathcal{R}$.

4 安全分析

4.1 正确性证明

首先, 针对方案各阶段的密文数据完整性验证, 已给出其正确性证明.

公式 (5) $\tilde{e}\left(g,\prod_{j=1}^{\beta}(\sigma_{i_j,2})^{\varsigma_{i_j}}\right)=\tilde{e}\left(y,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}}))^{\varsigma_{i_j}}\right)\prod_{j=1}^{\beta}\tilde{e}\left(\sigma_{i_j,1},(H_2(c_{i_j}\|t_{i_j}))^{\varsigma_{i_j}}\right)$ 的正确性推导过程如下:

$$\begin{aligned}\tilde{e}\left(g,\prod_{j=1}^{\beta}(\sigma_{i_j,2})^{\varsigma_{i_j}}\right)&=\tilde{e}\left(g,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}})^{\delta}\cdot H_2(c_{i_j}\|t_{i_j})^{w_{i_j}})^{\varsigma_{i_j}}\right)\\&=\tilde{e}\left(g,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}})^{\delta})^{\varsigma_{i_j}}\right)\tilde{e}\left(g^{w_{i_j}},\prod_{j=1}^{\beta}(H_2(c_{i_j}\|t_{i_j}))^{\varsigma_{i_j}}\right)\\&=\tilde{e}\left(y,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}}))^{\varsigma_{i_j}}\right)\prod_{j=1}^{\beta}\tilde{e}\left(\sigma_{i_j,1},(H_2(c_{i_j}\|t_{i_j}))^{\varsigma_{i_j}}\right).\end{aligned}$$

公式 (7) $\tilde{e}\left(g,\prod_{i=1}^{n}(\sigma_{i,2})^{\varsigma_i}\right)=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\sigma_{i,1},(H_2(c_i\|t_i))^{\varsigma_i})$ 的正确性推导过程如下:

$$\begin{aligned}\tilde{e}\left(g,\prod_{i=1}^{n}(\sigma_{i,2})^{\varsigma_i}\right)&=\tilde{e}\left(g,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i})^{\delta}\cdot H_2(c_i\|t_i)^{w_i})^{\varsigma_i}\right)\\&=\tilde{e}\left(g,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i})^{\varsigma_i})^{\delta}\right)\tilde{e}\left(g,\prod_{i=1}^{n}(H_2(c_i\|t_i)^{\varsigma_i})^{w_i}\right)\\&=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\sigma_{i,1},(H_2(c_i\|t_i))^{\varsigma_i}).\end{aligned}$$

公式 (9) $\tilde{e}\left(g,\prod_{i=1}^{n}(\tilde{\sigma}_{i,2})^{\varsigma_i}\right)=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\tilde{\sigma}_{i,1},(H_2(ct_i\|c_i\|V_i\|t_i))^{\varsigma_i})$ 的正确性推导过程如下:

$$\begin{aligned}\tilde{e}\left(g,\prod_{i=1}^{n}(\tilde{\sigma}_{i,2})^{\varsigma_i}\right)&=\tilde{e}\left(g,\prod_{i=1}^{n}(sk_i\cdot H_2(ct_i\|c_i\|V_i\|t_i)^{w_i})^{\varsigma_i}\right)\\&=\tilde{e}\left(g,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i})^{\delta})^{\varsigma_i}\right)\tilde{e}\left(g^{w_i},\prod_{i=1}^{n}(H_2(ct_i\|c_i\|V_i\|t_i))^{\varsigma_i}\right)\\&=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{\varsigma_i}\right)\prod_{i=1}^{n}\tilde{e}(\tilde{\sigma}_{i,1},(H_2(ct_i\|c_i\|V_i\|t_i))^{\varsigma_i}).\end{aligned}$$

公式 (13) $\tilde{e}(g,\sigma_{\mathcal{R},2})=\tilde{e}(y,H_1(id_{\mathcal{CS}}))\tilde{e}(\sigma_{\mathcal{R},1},H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}}))$ 的正确性推导过程如下:

$$\begin{aligned}\tilde{e}(g,\sigma_{\mathcal{R},2})&=\tilde{e}\left(g,sk_{\mathcal{CS}}\cdot H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})^{\varepsilon}\right)=\tilde{e}\left(g,H_1(id_{\mathcal{CS}})^{\delta}\right)\tilde{e}\left(g,H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})^{\varepsilon}\right)\\&=\tilde{e}(y,H_1(id_{\mathcal{CS}}))\tilde{e}(\sigma_{\mathcal{R},1},H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})).\end{aligned}$$

公式 (6) $\tilde{e}\left(g,\sigma_{i_j,2}\right)=\tilde{e}\left(y,H_1(id_{\mathcal{DO}_{i_j}})\right)\tilde{e}\left(\sigma_{i_j,1},H_2(c_{i_j}\|t_{i_j})\right)$ 的正确性推导类似于公式 (5), 同样, 公式 (8)、(10)–(12) 的正确性推导分别类似于公式 (7)、(9) 和 (13), 证明过程不再赘述.

4.2 安全性分析

定理 1. 基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案可确保数据拥有者数据在各阶段的机密性.

证明: $\mathcal{DO}_{i_j}$ 通过 ODMT-BGN 算法加密医疗数据 m_{i_j} 得到密文数据 $c_{i_j}=u_1^{m_{i_j}}u_2^{m_{i_j}^2}u_3^{r_{i_j}}$, 在 $\mathcal{ES}_i$ 对可验证信息 ψ_{i_j} 批

量验证通过后, $\mathcal{ES}_i$ 对所有密文数据 c_{i_j} 进行聚合得到边级聚合密文 $c_i = u_1^{\sum_{j=1}^{\beta} m_{i_j}} u_2^{\sum_{j=1}^{\beta} m_{i_j}^2} u_3^{\sum_{j=1}^{\beta} r_{i_j}}$. 本质上, c_i 是 $\sum_{j=1}^{\beta} m_{i_j}$ 和 $\sum_{j=1}^{\beta} m_{i_j}^2$ 的 ODMT-BGN 算法加密的密文. 当 $\mathcal{CS}$ 通过对 ψ_i 批量验证后, 再根据 $\mathcal{RC}$ 发起的分析请求 $\mathcal{Q}$, $\mathcal{CS}$ 对 ψ_i 中的密文 c_i 生成聚合密文 $C = \prod_{\mathcal{Q}\subseteq S} c_{\mathcal{Q}} = u_1^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta} m_{\mathcal{Q}_j}} u_2^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta} m_{\mathcal{Q}_j}^2} u_3^{\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta} r_{\mathcal{Q}_j}}$, $\mathcal{Q}\subseteq S, j\in[1,\beta]$. 本质上, C 是 $\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta} m_{\mathcal{Q}_j}$ 和 $\sum_{\mathcal{Q}\subseteq S}\sum_{j=1}^{\beta} m_{\mathcal{Q}_j}^2$ 也是 ODMT-BGN 算法加密后的密文. 而授权信息、分析请求以及分析结果的加密后的密文是一种基于身份的加密.

通过文献 [17] 安全分析可知: ODMT-BGN 沿用了 BGN 的安全假设, 即都归约为子群判定问题, 并且都能够通过抵抗选择明文攻击以保障其安全性. 即便敌手截获了任何阶段的密文, 也无法破解出明文信息. 此外, 通过文献 [30] 安全分析得知: 授权信息、分析请求与分析结果的加密满足 DBDH 假设, 同样保证授权信息、分析请求和分析结果在选择明文攻击下的不可区分性.

定理 2. 当 ODMT-BGN 算法与文献 [42] 中新型 BGN 同态加密的安全参数相同时, ODMT-BGN 算法群的阶数是文献 [42] 的 $1/2^{\lambda}$.

证明: 由 $\begin{cases} m < p_0' p_3' \\ m^2 < p_1' p_3' \end{cases} \Rightarrow \begin{cases} m < p_0' p_3' \\ m \leqslant \left\lfloor \sqrt{p_1' p_3'} \right\rfloor \end{cases}$ 可知, 当安全参数为 λ 时, 文献 [42] 的安全与本文 ODMT-BGN 算法安全性相同. 假设方案 [42] 中密文空间为 $\tilde{G}_1$, 因此得出本文密文空间的阶数与文献 [42] 密文空间的阶数之比为 $\dfrac{|G_1|}{|\tilde{G}_1|} = \dfrac{2^{\frac{\lambda}{2}}\times 2^{\frac{3\lambda}{2}}\times 2^{\frac{\lambda}{2}}\times 2^{\frac{\lambda}{2}}}{2^{\lambda}\times 2^{\lambda}\times 2^{\lambda}\times 2^{\lambda}} = \dfrac{1}{2^{\lambda}}$.

由定理 2 可知, ODMT-BGN 算法密文空间长度是文献 [42] 密文空间长度的 $\log_2 \dfrac{2^{4\lambda-\lambda}}{2^{4\lambda}} = \dfrac{3}{4}$.

定理 3. 基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案可确保各阶段数据的完整性.

证明: 基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案在数据传输过程中设计一个基于身份标识的签名来确保密文的完整性. 在密文数据上传阶段, $\mathcal{DO}_{i_j}$ 利用私钥 sk_{i_j} 对密文进行数字签名 $\sigma_{i_j} = (\sigma_{i_j,1}, \sigma_{i_j,2}) = \left(g^{w_{i_j}}, sk_{i_j}\cdot H_2(c_{i_j}\|t_{i_j})^{w_{i_j}}\right)$, 在边聚合验证与密文数据上传阶段, $\mathcal{ES}_i$ 利用私钥 sk_i 分别对辖区内的聚合密文 c_i 以及 $\mathcal{ES}_i$ 所属的密文数据标签 i 进行数字签名 $\sigma_i = (\sigma_{i,1},\sigma_{i,2}) = (g^{w_i}, sk_i\cdot H_2(c_i\|t_i)^{w_i})$ 和 $\tilde{\sigma}_i = (\tilde{\sigma}_{i,1},\tilde{\sigma}_{i,2}) = (g^{w_i}, sk_i\cdot H_2(ct_i\|c_i\|V_i\|t_i)^{w_i})$. 在统计分析阶段, $\mathcal{RC}$ 分别对授权信息 $\mathcal{I}$ 和分析请求 $\mathcal{Q}$ 进行数字签名 $\hat{\sigma}_i = (\hat{\sigma}_{i,1},\hat{\sigma}_{i,2}) = (g^{v_i}, sk_{\mathcal{RC}}\cdot H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}\|\hat{t}_i)^{v_i})$ 和 $\sigma_{\mathcal{Q}} = (\sigma_{\mathcal{Q},1},\sigma_{\mathcal{Q},2}) = (g^{v}, sk_{\mathcal{RC}}\cdot H_2(\mathcal{C}_{\mathcal{Q}}\|t_{\mathcal{Q}}\|\mathcal{AM})^{v})$, 以及 TEE 利用私钥 $sk_{\mathcal{CS}}$ 对密文统计分析结果进行数字签名 $\sigma_{\mathcal{R}} = (\sigma_{\mathcal{R},1},\sigma_{\mathcal{R},2}) = \left(g^{e}, sk_{\mathcal{CS}}\cdot H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})^{e}\right)$. 本方案采用的基于身份聚合算法, 沿用了文献 [30,43] 中的构造思路, 并且这两篇文献已经将该算法的不可伪造性安全性证明归约到 CDH 难题, 因此本方案的聚合签名依然具备此特性.

本方案分析计算过程是在可信执行环境 TEE (Intel SGX) 内进行, 其正确性依托硬件安全的保证, 更详细证明过程见文献 [27–29]. 因此, 只要各阶段密文在传输和聚合时总保持完整性, 敌手就无法通过替换或篡改数据来欺骗系统. 对此, 通过构建以下多个安全游戏来达成完整性保护目的, 即安全游戏 $Game_1 - Game_5$.

$Game_1$: 在密文数据上传阶段, 不同于方案的 ψ_{i_j}, 假设至少有一个数据拥有者 $id_{\mathcal{DO}_{i_\theta}}$ 的可验证数据信息 $\psi_{i_\theta} = \{c_{i_\theta},\sigma_{i_\theta},t_{i_\theta},id_{\mathcal{DO}_{i_\theta}}\}$ 在上传到 $\mathcal{ES}_i$ 的过程中, 其密文数据 c_{i_θ} 被敌手 $\mathcal{A}_1$ 以不可忽略的优势篡改或替换为 $c_{i_\theta}^*$, 并通过以下批量验证等式:

$$\tilde{e}\left(g,\prod_{j=1}^{\beta}(\sigma_{i_j,2})^{c_{i_j}}\right) = \tilde{e}\left(y,\prod_{j=1}^{\beta}(H_1(id_{\mathcal{DO}_{i_j}}))^{c_{i_j}}\right)\prod_{j\in[1,\beta]/\vartheta}\tilde{e}\left(\sigma_{i_j,1},(H_2(c_{i_j}\|t_{i_j}))^{c_{i_j}}\right)\cdot\tilde{e}\left(\sigma_{i_\theta,1},(H_2(c_{i_\theta}^*\|t_{i_\theta}))^{c_{i_\theta}}\right) \tag{14}$$

对于不同 $\mathcal{DO}_{i_j}$ 上传的可验证数据信息 ψ_{i_j} 应满足批量验证公式 (5), 根据公式 (5) 和 (14) 可以得到以下关系: $(H_2(c_{i_\theta}^*\|t_{i_\theta}))^{c_{i_\theta}} = (H_2(c_{i_\theta}\|t_{i_\theta}))^{c_{i_j}}$, 且设置 $\varpi = (H_2(c_{i_\theta}\|t_{i_\theta}))^{c_{i_j}}$. 若敌手 $\mathcal{A}_1$ 在多项式时间内通过替换或篡改等手段依旧实现完整性校验, 则意味着 $\mathcal{A}_1$ 能在多项式时间内解决 $H_2(c_{i_\theta}\|t_{i_\theta})$ 与 ϖ 的离散对数问题, 并找到 $H_2(c_{i_\theta}\|t_{i_\theta})$ 的哈希碰撞 $H_2(c_{i_\theta}^*\|t_{i_\theta})$. 这一假设既与离散对数问题的难以求解性相悖, 也与哈希函数 H_2 的抗碰撞属性相违背. 由此可知, $\mathcal{A}_1$ 无法以任何不可忽略优势在批量验证中胜出 $Game_1$, 从计算角度是不可行的.

$Game_2$: 在边聚合验证与密文数据上传阶段, c_i 是 $\mathcal{ES}_i$ 聚合后密文, 不同于方案正确生成可验证信息 $\psi_i=\{c_i,\sigma_i,ct_i,\tilde{\sigma}_i,V_i,id_{\mathcal{ES}_i},t_i\}$, 也假设至少有一个边缘服务器 $id_{\mathcal{ES}_\vartheta}$ 的 $\psi_\vartheta=\{c_\vartheta,\sigma_\vartheta,t_\vartheta,ct_\vartheta,\tilde{\sigma}_\vartheta,V_\vartheta\}$ 在上传到 $\mathcal{CS}$ 的过程中, 其密文数据 c_ϑ 被敌手 $\mathcal{A}_2$ 以不可忽略的优势篡改或替换为 c_ϑ^*, 并通过以下批量验证等式:

$$\tilde{e}\left(g,\prod_{i=1}^{n}(\sigma_{i,2})^{s_i}\right)=\tilde{e}\left(y,\prod_{i=1}^{n}(H_1(id_{\mathcal{ES}_i}))^{s_i}\right)\prod_{i\in[1,n]/\vartheta}\tilde{e}(\sigma_{i,1},(H_2(c_i\|t_i))^{s_i})\cdot\tilde{e}(\sigma_{\vartheta,1},(H_2(c_\vartheta^*\|t_\vartheta))^{s_i}) \tag{15}$$

对于每个 $\mathcal{ES}_i$ 上传的可验证数据信息 ψ_i 应满足批量验证公式 (7), 根据公式 (7) 和 (15) 可以得到: $(H_2(c_\vartheta^*\|t_\vartheta))^{s_i}=(H_2(c_i\|t_i))^{s_i}$, 且设置 $\varpi'=(H_2(c_\vartheta\|t_\vartheta))^{s_i}$. 若敌手 $\mathcal{A}_2$ 在多项式时间内通过替换或篡改形式依旧能完成完整性校验, 则意味着 $\mathcal{A}_2$ 能在多项式时间内解决 $H_2(c_\vartheta\|t_\vartheta)$ 与 ϖ' 的离散对数问题, 并且可以找到 $H_2(c_\vartheta^*\|t_\vartheta)$ 的哈希碰撞 $H_2(c_i\|t_i)$. 这一假设既与离散对数问题的难以求解性相悖, 也与哈希函数 H_2 的抗碰撞属性相违背. 由此可知, $\mathcal{A}_2$ 无法以任何不可忽略优势在批量验证中赢得 $Game_2$, 从计算角度是不可行的. 同理, 针对密文数据标签 ct_ϑ 来说, $\mathcal{A}_2$ 以不可忽略的优势赢得 $Game_2$ 在计算上也是不可行的.

$Game_3$: 在授权阶段, 不同于 $\mathcal{RC}$ 发送的可验证数据信息 $\hat{\psi}_i=\{\mathcal{CI}_i,\hat{\sigma}_i,\hat{t}_i,id_{\mathcal{RC}}\}$, 同样假设敌手 $\mathcal{A}_3$ 能够以不可忽略优势将 $\mathcal{CI}_{i,2}$ 替换或篡改为 $\mathcal{CI}_{i,2}^*$ 以通过下面的验证:

$$\tilde{e}(g,\hat{\sigma}_{i,2})=\tilde{e}(y,H_1(id_{\mathcal{RC}}))\tilde{e}\left(\hat{\sigma}_{i,1},H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}^*\|\hat{t}_i)\right) \tag{16}$$

根据公式 (11) 和 (16) 可得: 若敌手 $\mathcal{A}_3$ 在多项式时间内通过替换或篡改形式依旧能完成完整性校验, 则意味着 $\mathcal{A}_3$ 能在多项式时间内解决 $H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}\|\hat{t}_i)$ 与 $H_2(\mathcal{CI}_{i,1}\|\mathcal{CI}_{i,2}^*\|\hat{t}_i)$ 的哈希碰撞问题, 这一假设与哈希函数 H_2 的抗碰撞属性相违背. 由此可知, $\mathcal{A}_3$ 无法以任何不可忽略优势在分析请求验证中赢得 $Game_3$, 从计算角度是不可行的.

$Game_4$: 在统计分析阶段, 不同于 $\mathcal{RC}$ 发送的可验证数据信息 $\psi_{\mathcal{Q}}=\{C_{\mathcal{Q}},\sigma_{\mathcal{Q}},t_{\mathcal{RC}},id_{\mathcal{RC}},\mathcal{AM}\}$, 同样假设敌手 $\mathcal{A}_4$ 能够以不可忽略优势将 $\mathcal{C}_{\mathcal{Q}}$ 替换或篡改为 $\mathcal{C}_{\mathcal{Q}}^*$ 以通过下列的验证:

$$\tilde{e}(g,\sigma_{\mathcal{Q},2})=\tilde{e}(y,H_1(id_{\mathcal{RC}}))\tilde{e}\left(\sigma_{\mathcal{Q},1},H_2(\mathcal{C}_{\mathcal{Q}}^*\|t_{\mathcal{Q}}\|\mathcal{AM})\right) \tag{17}$$

根据公式 (12) 和 (17) 可得: $H_2(\mathcal{C}_{\mathcal{Q}}\|t_{\mathcal{Q}}\|\mathcal{AM})=H_2(\mathcal{C}_{\mathcal{Q}}^*\|t_{\mathcal{Q}}\|\mathcal{AM})$. 若敌手 $\mathcal{A}_4$ 在多项式时间内通过替换或篡改形式依旧能完成完整性校验, 则意味着 $\mathcal{A}_4$ 能在多项式时间内解决 $H_2(\mathcal{C}_{\mathcal{Q}}\|t_{\mathcal{Q}}\|\mathcal{AM})$ 与 $H_2(\mathcal{C}_{\mathcal{Q}}^*\|t_{\mathcal{Q}}\|\mathcal{AM})$ 的哈希碰撞问题, 这一假设与哈希函数 H_2 的抗碰撞属性相违背. 由此可知, $\mathcal{A}_4$ 无法以任何不可忽略优势在分析请求验证中赢得 $Game_4$, 从计算角度是不可行的.

$Game_5$: 在验证与解密阶段, 不同于 $\mathcal{CS}$ 发送的可验证数据信息 $\psi=\{C_{\mathcal{R}},\sigma_{\mathcal{R}},t_{\mathcal{R}},id_{\mathcal{CS}}\}$, 同样假设敌手 $\mathcal{A}_5$ 能够以不可忽略优势将 $c_{\mathcal{R},2}$ 替换或篡改为 $c_{\mathcal{R},2}^*$ 以通过下列的验证:

$$\tilde{e}(g,\sigma_{\mathcal{R},2})=\tilde{e}(y,H_1(id_{\mathcal{CS}}))\tilde{e}\left(\sigma_{\mathcal{R},1},H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}^*\|t_{\mathcal{R}})\right) \tag{18}$$

由公式 (13) 和 (18) 可得: $H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})=H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}^*\|t_{\mathcal{R}})$. 若敌手 $\mathcal{A}_5$ 在多项式时间内通过替换或篡改形式依旧能完成完整性校验, 则意味着 $\mathcal{A}_5$ 能在多项式时间内解决 $H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}\|t_{\mathcal{R}})$ 与 $H_2(c_{\mathcal{R},1}\|c_{\mathcal{R},2}^*\|t_{\mathcal{R}})$ 的哈希碰撞问题, 这一假设与哈希函数 H_2 的抗碰撞属性相违背. 由此可知, $\mathcal{A}_5$ 无法以任何不可忽略优势在分析结果验证中赢得 $Game_5$, 从计算角度上是不可行的.

综上, 根据 $Game_1-Game_5$ 的安全分析过程得出, 基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案可确保各个阶段密文数据完整性. □

5 实验分析

5.1 安全性与功能性比较

首先, 将本文方案与相关方案[20,23,25,30,42]从安全和功能两方面进行对比分析, 对比结果如表 1 所示. 本方案同时具备数据多源、数据的机密性、完整性、数据访问控制支持方差、期望细粒度统计分析、支持双消息形式加密以及硬件支持. 同时, 该方案全程保障了密文在传输与存储环节的机密性与完整性.

表 1　安全性与功能性比较

方案	数据多源	机密性	完整性验证	访问控制	细粒度分析	双消息形式	Intel SGX
文献[20]	√	√	√	√	×	×	×
文献[23]	√	√	√	√	×	×	×
文献[25]	√	√	×	×	×	×	×
文献[30]	√	√	√	×	×	×	×
文献[42]	√	√	√	×	√	√	×
本方案	√	√	√	√	√	√	√

5.2　计算与通信开销比较

本方案采用的硬件为: Intel(R) Core(TM) i7-8565U CPU @ 1.80 GHz 和 8 GB 内存条与 Intel(R) Core(TM) i7-10510U CPU @ (1.8 GHz × 8), 32 GB 内存, Ubuntu 20.04.3 LTS 操作系统, SGX 开发工具包为 sgx_linux_x64_sdk_2.19. 我们基于 BRFSS 数据集 (https://www.kaggle.com/health-dataset=hypertension_data.csv), 收集了美国 50 个州及哥伦比亚特区每年超过 40 万名居民的慢性病相关数据. 方案的实现依托密码学库的基础数学和配对操作, 采用 C 和 Java 语言开发完成. 同时, 我们采用密码学库中定义椭圆曲线实现本协议, 其中 p_0、p_2 和 p_3 都是 256 位的大素数, 而 p_1 则是 3×256 位大素数, q 是 512 位大素数. 由定理 2 可知, 我们设置参数可以与文献 [42] 达到相同安全条件. 设定 P 是一次双线性配对操作, E 是一次模指数运算, $\mathcal{H}$ 是一次将消息映射为椭圆曲线中某坐标点的运算, $\mathcal{M}u$ 表示在椭圆曲线上执行一次倍点运算, Mu 表示执行一次模乘法运算, h 表示执行一次哈希运算, Ad 表示在基于椭圆曲线的加法循环群中执行一次加法运算. 其中, β 代表每个边缘服务器含有数据数量, n 代表边缘服务器的数量, s 代表请求访问边缘服务器的数量且 $1\leqslant s\leqslant n$.

在计算开销方面, 我们将本方案与各类聚合方案[20,23,30,42]在数据拥有者侧、边缘服务器侧以及统计分析侧这 3 方面进行计算开销对比, 对比结果如表 2 所示. 其中, 本方案将云服务器与研究中心结合共同执行统计分析功能与其他方案模型中实体执行统计分析功能相比较. 由表 2 可知, 统计分析过程总共不仅涉及 $s+10$ 次模指数运算与 7 次哈希运算, 还需要执行 8 次双线性配对和 $3s+4$ 次模乘, 故此统计分析侧的总开销为 $(s+10)E+7\mathcal{H}+8P+(3s+4)Mu$. 这是由于本方案的数据统计分析过程是在 $\mathcal{CS}$ 的 TEE 内部进行的明文数据统计, 研究中心仅需要一次验证和解密操作即可获得统计分析结果, 文献 [20,23] 需要在云服务器上基于密文进行计算与验证, 而文献 [42] 则是将计算的中间变量的密文发送至研究中心, 研究中心需要进行一次验证和 2 次 BGN 同态解密算法得到中间结果, 再利用中间结果进行统计分析的计算. 因此, 本方案的统计分析效率高于文献 [20,23,42].

表 2　计算开销比较

方案	数据拥有者侧	边缘服务器侧	统计分析侧
文献[20]	$3E+\mathcal{M}u+2\mathcal{H}+2P+Mu$	$\mathcal{M}u+(2\beta+1)\mathcal{H}+2(\beta+1)P+2(\beta-1)Mu+2(\beta-1)Ad$	$nE+n\mathcal{H}+2nP+nMu$
文献[23]	$4E+\mathcal{H}+2Mu$	$2E+(\beta+1)\mathcal{H}+(\beta+1)P+(3\beta-2)Mu$	$E+n\mathcal{H}+(n+1)P+3(n-1)Mu$
文献[30]	$4E+\mathcal{H}+2Mu$	$(3\beta+2)E+(2\beta+1)\mathcal{H}+(\beta+2)P+(4\beta-3)Mu$	$E+2\mathcal{H}+4P$
文献[42]	$6E+\mathcal{H}+h+4Mu$	$(2\beta+3)E+2\mathcal{H}+(\beta+1)h+2P+(6\beta-1)Mu$	$(n+5)E+3\mathcal{H}+(n+2)h+6P+(7n-3)Mu$
本方案	$5E+\mathcal{H}+3Mu$	$(3\beta+11)E+(2\beta+5)\mathcal{H}+(\beta+6)P+(4\beta+3)Mu$	$(s+10)E+7\mathcal{H}+8P+(3s+4)Mu$

在通信开销方面, 由初始化算法可知, 将循环群 $\mathcal{G}_3$ 与循环群 $\mathcal{G}_1$ 都设置大小一样的 q 阶循环群, 同时也可以设置成大小不同循环群. 然而, 在实际部署过程一般都是将循环群 $\mathcal{G}_3$ 与循环群 $\mathcal{G}_1$ 设置成相同的群. 本方案通信部分主要包含数据拥有者上传密文阶段、边缘服务器上传聚合密文阶段及授权访问阶段这 3 个阶段通信开销对比. 由于本文选取 256 位大素数与文献 [42] 达到相同的安全程度, 且在明文取值范围一致前提下, 方案的密文会缩小到原来的 3/4, 所以在数据拥有者侧通信开销小于文献 [42]. 而边缘侧通信开销会比文献 [42] 略微高一些, 这是由于

引入访问控制机制所造成. 此外, 方案的授权访问通信开销小于文献 [20,23], 主要原因在于研究中心请求授权对象是边缘服务器, 而不是数据拥有者. 更为详细对比结果如表 3 所示.

表 3 通信开销比较

方案	数据拥有者侧	边缘侧	授权访问
文献[20]	$3\|\mathcal{G}_1\|$	$3(\beta+1)\|\mathcal{G}_1\|$	$2n(\beta+1)\|\mathcal{G}_1\|$
文献[23]	$2\|\mathcal{G}_1\|$	$2(\beta+1)\|\mathcal{G}_1\|$	$\beta n\|\mathcal{G}_1\|+(\beta+n)\|Z_q\|$
文献[30]	$2\|\mathcal{G}_1\|+\|\mathcal{G}_2\|$	$(2\beta+3)\|\mathcal{G}_1\|+(\beta+1)\|\mathcal{G}_2\|$	—
文献[42]	$\|G_1\|+2\|\mathcal{G}_1\|$	$(\beta+1)\|G_1\|+2(\beta+1)\|\mathcal{G}_1\|$	—
本方案	$\frac{3}{4}\|G_1\|+2\|\mathcal{G}_1\|$	$\frac{3}{4}(\beta+1)\|G_1\|+(2\beta+5)\|\mathcal{G}_1\|$	$6s\|\mathcal{G}_1\|+2s\|\mathcal{G}_2\|$

图 3 表示当边缘服务器数量 $n=50$ 时, 随着数据拥有者的数量增加, 方案中各类算法的计算开销. 其中, 数据拥有者数量与边聚合验证、边聚合密文以及初始化成正比, 而不影响边签名和云统计分析操作的计算开销. 图 4 表示当数据拥有者 $\beta=1000$ 时, 方案中各类算法计算开销. 其中, 边缘服务器数量与云聚合验证、云聚合密文以及初始化成正比, 而与云签名、云统计分析计算开销关系不大.

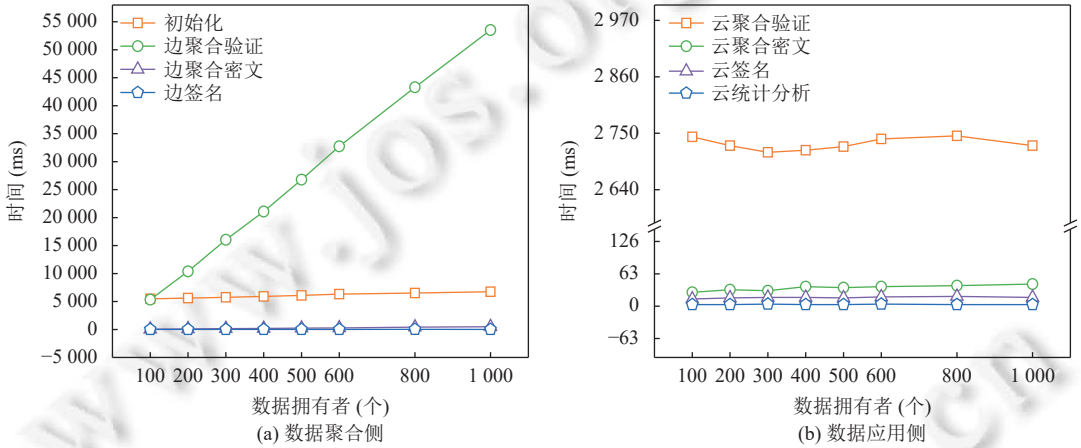


图 3 当 $n=50$ 时方案中各类算法开销

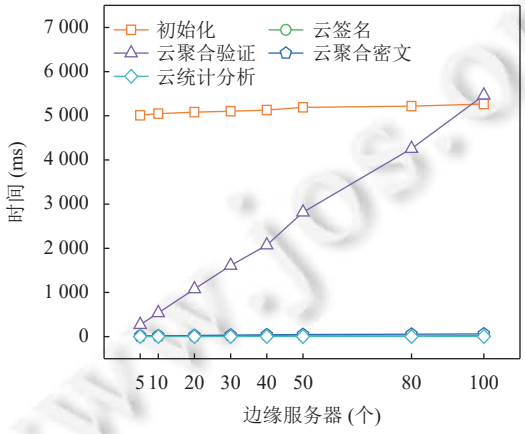


图 4 当 $\beta=1000$ 时方案各类算法开销

图 5(a) 表示当边缘服务器数量 $n=50$ 时, 随着每个边缘服务器辖区内数据拥有者数量的增加, 方案中加解密的计算开销对比, 图 5(b) 表示当数据拥有者 $\beta=1000$ 时, 随着边缘服务器数量的增加, 方案中加解密计算开销对

比. 由图 5(a) 和图 5(b) 可知, 不论数据拥有者数量增加还是边缘服务器数量增加, 都会造成云解密计算开销的增加. 其中图 5(a) 和图 5(b) 中云解密算法的计算开销为图右侧纵坐标.

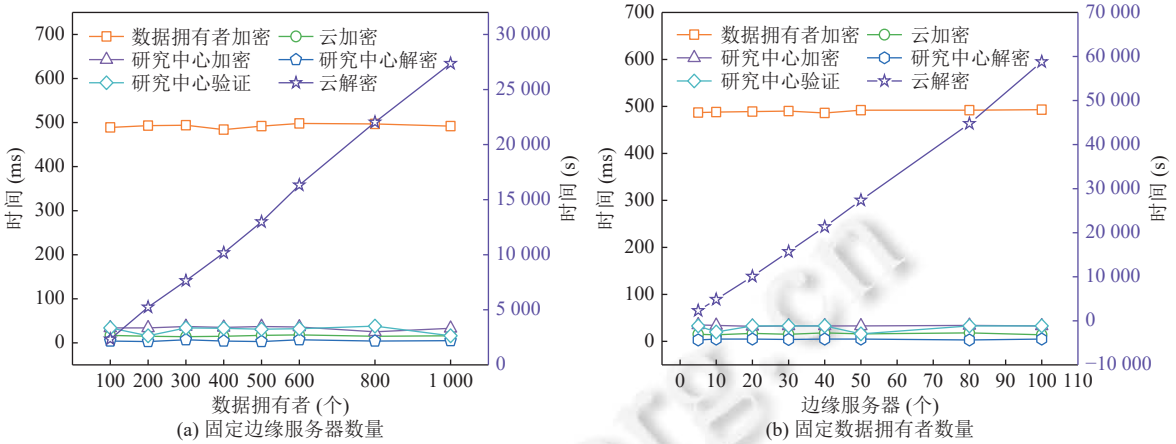


图 5 加解密算法的计算开销

图 6(a) 表示当边缘服务器数量 $n = 50$ 时, 随着每个边缘服务器辖区内数据拥有者数量的增加, 边缘服务器聚合验证开销越优于边缘服务器逐个验证的累计开销和. 图 6(b) 则表示当数据拥有者 $\beta = 1000$ 时, 随着边缘服务器数量的增加, 云服务器逐个验证的累计开销和始终大于云服务器聚合验证的计算开销.

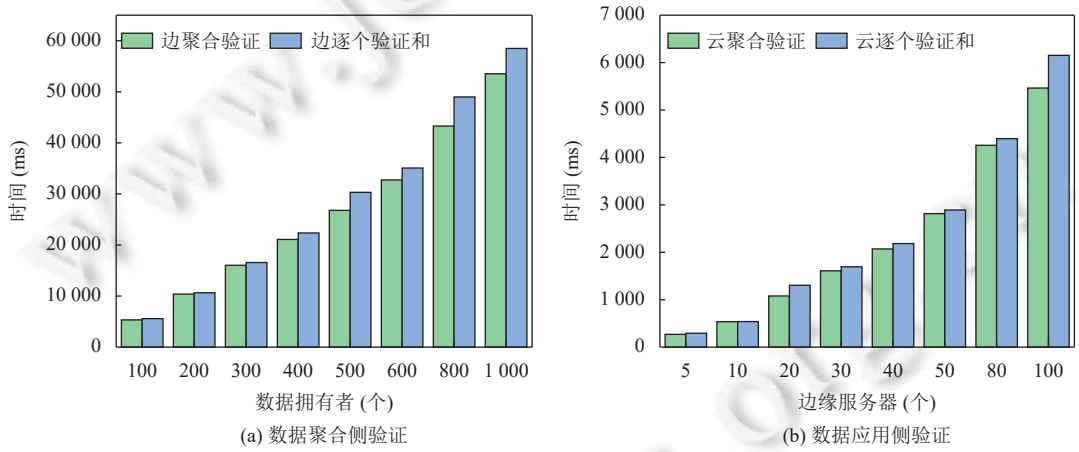
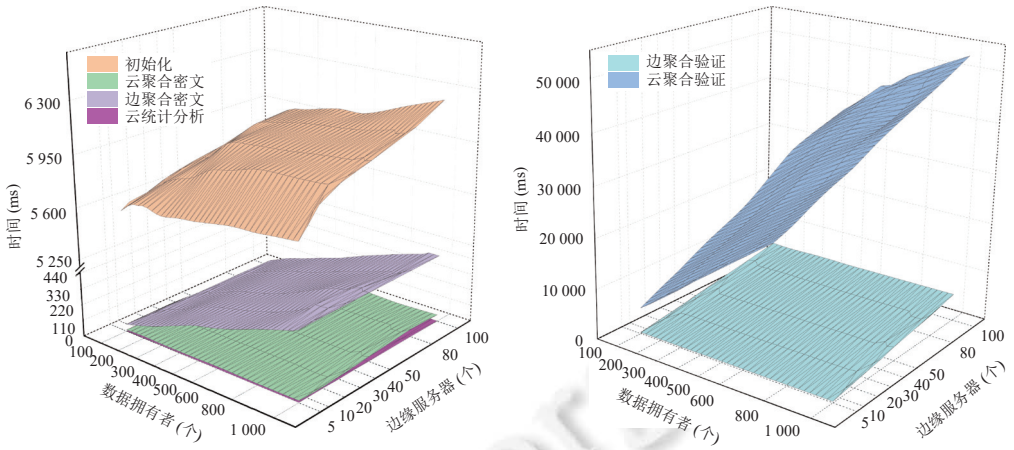


图 6 聚合验证与逐个验证累计和的计算开销对比

图 7 表示当边缘服务器数量与数据拥有者数量同时变化时, 云服务器的聚合验证和聚合密文的计算开销与边缘服务器数量有关, 而与数据拥有者数量无关. 而边缘服务聚合验证和聚合密文的计算开销只与数据拥有者数量有关, 而与边缘服务器数量无关. 此外, 由于初始化包含为数据拥有者和边缘服务器生成签名密钥的过程, 因此初始计算开销会随着二者数量一同变化, 然而云统计分析的计算开销与二者数量均无关系.

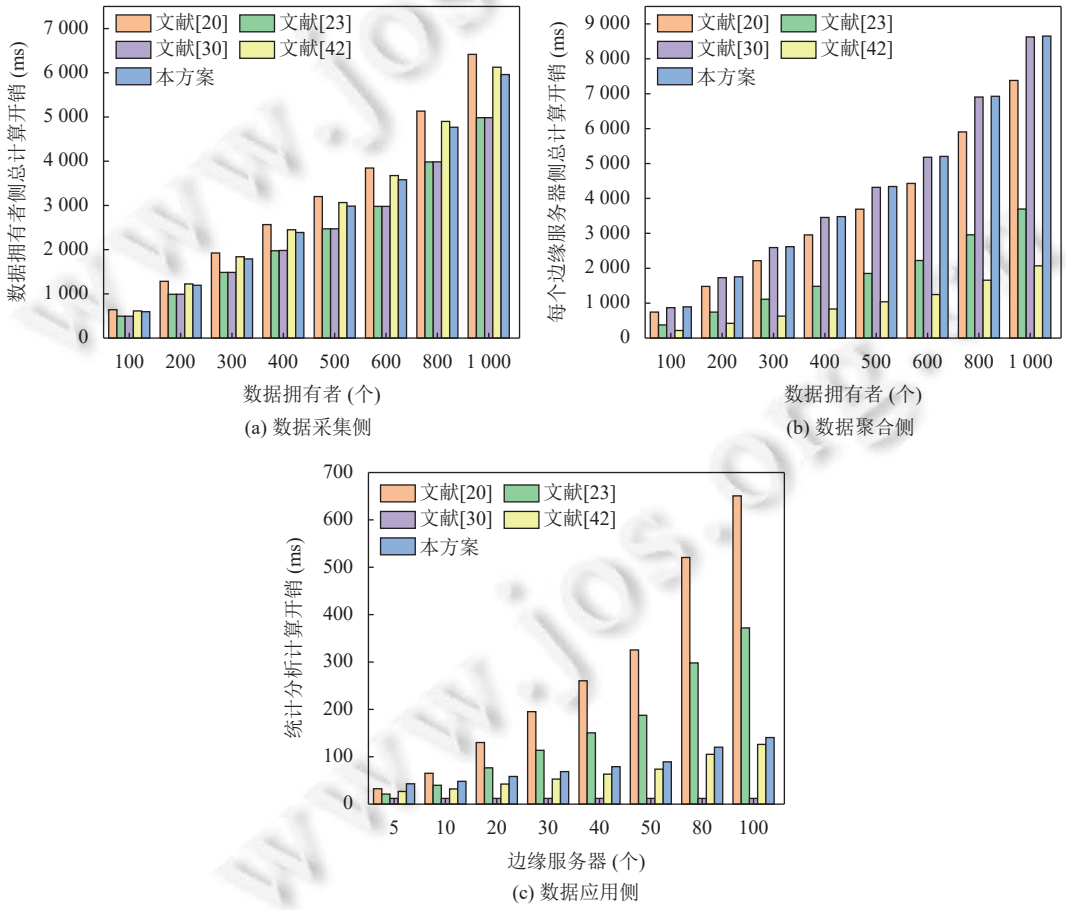
图 8 表示所提出的方案与其他研究工作在数据拥有者侧、边缘服务器侧和统计分析侧的计算开销对比. 图 8(a) 表明, 本方案在数据拥有者侧的计算开销高于文献 [23,30], 而低于文献 [20,42], 主要原因是在数据拥有者侧保持密文具有和、方差等特征, 而文献 [23,30] 均不满足. 图 8(b) 表明, 由于在边缘服务器侧引入授权机制导致该方案在边缘服务器侧总计算开销均高于其他文献. 图 8(c) 表明, 相比文献 [20,23] 其密文形式的计算开销来说, 此方案设计基于 TEE 明文统计分析计算开销更低. 由于研究中心增加了请求授权的过程, 图 8(c) 仅呈现出当 $s = n$ 条件下统计分析计算开销略高于文献 [42] 的情况, 而当 $s < n$ 进行细粒度分析时, 统计分析计算开销将会低于文

献 [42]. 此外, 文献 [30] 的统计分析仅是密文聚合后的累加值, 而不存在后续分析过程. 因此, 文献 [30] 所呈现效率最高.



(a) 初始化、云边聚合与分析 (b) 云边聚合验证

图 7 本方案计算开销



(a) 数据采集侧 (b) 数据聚合侧

(c) 数据应用侧

图 8 各个方案计算开销

6 总 结

本文设计并提出了一种基于 TEE 安全高效的细粒度统计分析与可验证数据聚合方案. 通过优化 BGN 同态加密的参数, 设计了一种 ODMT-BGN 算法, 此算法在相同大小的安全参数下能提供更高的安全强度并扩展了消息空间, 保障了数据的机密性; 其次, 本方案依托于 TEE 实现了多种明文数据的统计分析方法, 不仅保证了数据分析过程的安全性, 还明显提高了统计分析的效率. 此外, 研究中心仅进行一次解密就可以得到方差、期望等统计分析结果, 降低了研究中心的计算代价; 最后, 构建了一种依托身份的聚合签名与验证方法, 从而维护了数据传输和存储过程中的完整性, 也通过批量认证降低了认证成本, 并实现了医疗数据的访问控制, 满足了研究中心细粒度统计分析的需求. 特别地, 随着数据拥有者和边缘服务器数量的增加, 研究中心能够以近乎恒定的计算开销完成统计结果的验证和解密. 在后续研究中, 将进一步扩展该方案的分析方法种类, 优化 TEE 中的最小算子达到更高的计算安全性, 并探讨方案在去中心化场景中的应用可行性.

References

[1] Zhang YQ, Wang XF, Liu XF, Liu L. Survey on cloud computing security. Ruan Jian Xue Bao/Journal of Software, 2016, 27(6): 1328–1348 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5004.htm [doi: 10.13328/j.cnki.jos.005004]

[2] Feng DG, Zhang M, Zhang Y, Xu Z. Study on cloud computing security. Ruan Jian Xue Bao/Journal of Software, 2011, 22(1): 71–83 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/SP.J.1001.2011.03958]

[3] Botta A, De Donato W, Persico V, Pescapé A. Integration of cloud computing and Internet of Things: A survey. Future Generation Computer Systems, 2016, 56: 684–700. [doi: 10.1016/j.future.2015.09.021]

[4] Marinescu DC. Cloud Computing: Theory and Practice. 3rd ed., Amsterdam: Elsevier, 2022. [doi: 10.1016/C2020-0-02233-4]

[5] Catarinucci L, De Donno D, Mainetti L, Palano L, Patrono L, Stefanizzi ML, Tarricone L. An IoT-aware architecture for smart healthcare systems. IEEE Internet of Things Journal, 2015, 2(6): 515–526. [doi: 10.1109/JIOT.2015.2417684]

[6] Mohamad Noor MB, Hassan WH. Current research on Internet of Things (IoT) security: A survey. Computer Networks, 2019, 148: 283–294. [doi: 10.1016/j.comnet.2018.11.025]

[7] Bansal M, Nanda M, Husain MN. Security and privacy aspects for Internet of Things (IoT). In: Proc. of the 6th Int'l Conf. on Inventive Computation Technologies (ICICT). Coimbatore: IEEE, 2021. 199–204. [doi: 10.1109/ICICT50816.2021.9358665]

[8] Chen DY, Zhao H. Data security and privacy protection issues in cloud computing. In: Proc. of the 2012 Int'l Conf. on Computer Science and Electronics Engineering. Hangzhou: IEEE, 2012. 647–651. [doi: 10.1109/ICCSEE.2012.193]

[9] Choudhury T, Gupta A, Pradhan S, Kumar P, Rathore YS. Privacy and security of cloud-based Internet of Things (IoT). In: Proc. of the 3rd Int'l Conf. on Computational Intelligence and Networks (CINE). Odisha: IEEE, 2017. 40–45. [doi: 10.1109/CINE.2017.28]

[10] Qiu J, Tian ZH, Du CL, Zuo Q, Su S, Fang BX. A survey on access control in the age of Internet of Things. IEEE Internet of Things Journal, 2020, 7(6): 4682–4696. [doi: 10.1109/JIOT.2020.2969326]

[11] Goldwasser S, Micali S, Rivest RL. A digital signature scheme secure against adaptive chosen-message attacks. SIAM Journal on Computing, 1988, 17(2): 281–308. [doi: 10.1137/0217017]

[12] Xiong H, Bao YY, Nie XY, Asoor YI. Server-aided attribute-based signature supporting expressive access structures for industrial Internet of Things. IEEE Trans. on Industrial Informatics, 2020, 16(2): 1013–1023. [doi: 10.1109/TII.2019.2921516]

[13] Hu CQ, Pu YW, Yang FH, Zhao RF, Alrawais A, Xiang T. Secure and efficient data collection and storage of IoT in smart ocean. IEEE Internet of Things Journal, 2020, 7(10): 9980–9994. [doi: 10.1109/JIOT.2020.2988733]

[14] Ducas L, Kiltz E, Lepoint T, Lyubashevsky V, Schwabe P, Seiler G, Stehlé D. Crystals-Dilithium: A lattice-based digital signature scheme. IACR Trans. on Cryptographic Hardware and Embedded Systems, 2018, 2018(1): 238–268. [doi: 10.13154/tches.v2018.i1.238-268]

[15] Hamdi M, Pirbhulal S, Abie H. A homomorphic digital signature scheme for the Internet of Things. 2022. [doi: 10.20944/preprints202202.0238.v1]

[16] Zhao YQ, Yang XY, Feng Q, Yu Y. Anonymous credential protocol based on SM2 digital signature. Ruan Jian Xue Bao/Journal of Software, 2024, 35(7): 3469–3481 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6929.htm [doi: 10.13328/j.cnki.jos.006929]

[17] Boneh D, Goh EJ, Nissim K. Evaluating 2-DNF formulas on ciphertexts. In: Proc. of the 2nd Theory of Cryptography Conf. on Theory of Cryptography. Cambridge: Springer, 2005. 325–341. [doi: 10.1007/978-3-540-30576-7_18]

[18] Li FJ, Luo B, Liu P. Secure information aggregation for smart grids using homomorphic encryption. In: Proc. of the 1st IEEE Int'l Conf. on Smart Grid Communications. Gaithersburg: IEEE, 2010. 327–332. [doi: 10.1109/SMARTGRID.2010.5622064]

[19] Zeng ZX, Liu YN, Chang L. A robust and optional privacy data aggregation scheme for fog-enhanced IoT network. IEEE Systems Journal, 2023, 17(1): 1110–1120. [doi: 10.1109/JSYST.2022.3177418]

[20] Wang HQ, Wang ZW, Domingo-Ferrer J. Anonymous and secure aggregation scheme in fog-based public cloud computing. Future Generation Computer Systems, 2018, 78: 712–719. [doi: 10.1016/j.future.2017.02.032]

[21] Lu RX, Liang XH, Li X, Lin XD, Shen XM. EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. IEEE Trans. on Parallel and Distributed Systems, 2012, 23(9): 1621–1631. [doi: 10.1109/TPDS.2012.86]

[22] Wang XD, Garg S, Lin H, Kaddoum G, Hu J, Hossain MS. A secure data aggregation strategy in edge computing and blockchain-empowered Internet of Things. IEEE Internet of Things Journal, 2022, 9(16): 14237–14246. [doi: 10.1109/JIOT.2020.3023588]

[23] Li X, Liu SP, Wu F, Kumari S, Rodrigues JJPC. Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications. IEEE Internet of Things Journal, 2019, 6(3): 4755–4763. [doi: 10.1109/JIOT.2018.2874473]

[24] Fan CI, Huang SY, Lai YL. Privacy-enhanced data aggregation scheme against internal attackers in smart grid. IEEE Trans. on Industrial Informatics, 2014, 10(1): 666–675. [doi: 10.1109/TII.2013.2277938]

[25] Tang WJ, Ren J, Deng K, Zhang YX. Secure data aggregation of lightweight E-healthcare IoT devices with fair incentives. IEEE Internet of Things Journal, 2019, 6(5): 8714–8726. [doi: 10.1109/JIOT.2019.2923261]

[26] Han S, Zhao S, Li QH, Ju CH, Zhou WL. PPM-HDA: Privacy-preserving and multifunctional health data aggregation with fault tolerance. IEEE Trans. on Information Forensics and Security, 2016, 11(9): 1940–1955. [doi: 10.1109/TIFS.2015.2472369]

[27] McKeen F, Alexandrovich I, Berenzon A, Rozas C, Shafi H, Shanbhogue V, Savagaonkar UR. Innovative instructions and software model for isolated execution. In: Proc. of the 2nd Int'l Workshop on Hardware and Architectural Support for Security and Privacy. Tel-Aviv: ACM, 2013. 10. [doi: 10.1145/2487726.2488368]

[28] Costan V, Devadas S. Intel SGX explained. IACR Cryptology ePrint Archive, 2016. https://eprint.iacr.org/search?q=Intel+SGX+Explained

[29] Intel® Software Guard Extensions. Intel® software guard extensions SDK for Linux* OS. 2023. https://www.intel.com/content/www/us/en/developer/tools/software-guard-extensions/linux-overview.html

[30] Wang ZW. An identity-based data aggregation protocol for the smart grid. IEEE Trans. on Industrial Informatics, 2017, 13(5): 2428–2435. [doi: 10.1109/TII.2017.2705218]

[31] Cheng X, Chen FL, Xie D, Sun H, Huang C. Design of a secure medical data sharing scheme based on blockchain. Journal of Medical Systems, 2020, 44(2): 52. [doi: 10.1007/s10916-019-1468-1]

[32] Wang YT, Su Z, Xu QC, Li RD, Luan TH, Wang PH. A secure and intelligent data sharing scheme for UAV-assisted disaster rescue. IEEE/ACM Trans. on Networking, 2023, 31(6): 2422–2438. [doi: 10.1109/TNET.2022.3226458]

[33] Mollah MB, Azad MAK, Vasilakos A. Secure data sharing and searching at the edge of cloud-assisted Internet of Things. IEEE Cloud Computing, 2017, 4(1): 34–42. [doi: 10.1109/MCC.2017.9]

[34] Zhu XL, Hou HY, Fu SJ, Zhao YL, Liu B. Secure data sharing solution for portable health clinics. Ruan Jian Xue Bao/Journal of Software, 2023, 34(9): 4256–4274 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6638.htm [doi: 10.13328/j.cnki.jos.006638]

[35] Wang T, Wang JY, Yang QL, Yang B, Li HY, Xu F, Qiao ZR. An efficient verifiable searchable encryption scheme with aggregating authorization for blockchain-enabled IoT. IEEE Internet of Things Journal, 2022, 9(20): 20666–20680. [doi: 10.1109/JIOT.2022.3175859]

[36] Xu SM, Han XS, Xu GW, Ning JT, Huang XY, Deng RH. An adaptive secure and practical data sharing system with verifiable outsourced decryption. IEEE Trans. on Services Computing, 2024, 17(3): 776–788. [doi: 10.1109/TSC.2023.3321314]

[37] Liu ZL, Li T, Li P, Jia CF, Li J. Verifiable searchable encryption with aggregate keys for data sharing system. Future Generation Computer Systems, 2018, 78: 778–788. [doi: 10.1016/j.future.2017.02.024]

[38] Niu SF, Song M, Fang LZ, Wang CF. Cloud storage data sharing based on attribute encryption in smart healthcare. Journal of Electronics & Information Technology, 2022, 44(1): 107–117 (in Chinese with English abstract). [doi: 10.11999/JEIT210858]

[39] Boneh D, Franklin M. Identity-based encryption from the Weil pairing. In: Proc. of the 21st Annual Int'l Cryptology Conf. on Advances in Cryptology. Santa Barbara: Springer, 2001. 213–229. [doi: 10.1007/3-540-44647-8_13]

[40] Schueffel P, Groeneweg N, Baldegger R. The Crypto Encyclopedia: Coins, Tokens and Digital Assets from A to Z. Bern: Growth Publisher, 2019.

[41] Katz J, Lindell Y. Introduction to Modern Cryptography: Principles and Protocols. New York: Chapman and Hall/CRC, 2007. [doi: 10.

1201/9781420010756]

[42] Zhang XJ, Zhang JW, Huang C, Gu DW, Zhang Y. Verifiable encrypted medical data aggregation and statistical analysis scheme. Ruan Jian Xue Bao/Journal of Software, 2022, 33(11): 4285–4304 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6343.htm [doi: 10.13328/j.cnki.jos.006343]

[43] Gentry C, Ramzan Z. Identity-based aggregate signatures. In: Proc. of the 9th Int'l Conf. on Theory and Practice in Public Key Cryptography. New York: Springer, 2006. 257–273. [doi: 10.1007/11745853_17]

附中文参考文献

[1] 张玉清, 王晓菲, 刘雪峰, 刘玲. 云计算环境安全综述. 软件学报, 2016, 27(6): 1328–1348. http://www.jos.org.cn/1000-9825/5004.htm [doi: 10.13328/j.cnki.jos.005004]

[2] 冯登国, 张敏, 张妍, 徐震. 云计算安全研究. 软件学报, 2011, 22(1): 71–83. http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/SP.J.1001.2011.03958]

[16] 赵艳琦, 杨晓艺, 冯琦, 禹勇. 基于 SM2 数字签名的匿名凭证协议. 软件学报, 2024, 35(7): 3469–3481. http://www.jos.org.cn/1000-9825/6929.htm [doi: 10.13328/j.cnki.jos.006929]

[34] 朱雪岭, 侯慧莹, 付绍静, 赵运磊, 刘波. 面向便携式诊所的安全数据共享方案. 软件学报, 2023, 34(9): 4256–4274. http://www.jos.org.cn/1000-9825/6638.htm [doi: 10.13328/j.cnki.jos.006638]

[38] 牛淑芬, 宋蜜, 方丽芝, 王彩芬. 智慧医疗中基于属性加密的云存储数据共享. 电子与信息学报, 2022, 44(1): 107–117. [doi: 10.11999/JEIT210858]

[42] 张晓均, 张经伟, 黄超, 谷大武, 张源. 可验证医疗密态数据聚合与统计分析方案. 软件学报, 2022, 33(11): 4285–4304. http://www.jos.org.cn/1000-9825/6343.htm [doi: 10.13328/j.cnki.jos.006343]

作者简介

李鲍, 博士, 主要研究领域为隐私计算, 可信计算, 数据安全.

周福才, 博士, 教授, 博士生导师, 主要研究领域为隐私保护, 网络与信息安全, 云存储安全与认证技术, 大数据管理与隐私保护.

王强, 博士, 副教授, CCF 专业会员, 主要研究领域为隐私计算, 区块链技术及其应用, 云计算安全, 数据安全.

冯达, 博士, 主要研究领域为安全外包计算, 应用密码学, 数据安全.