·热点与综述· 文章编号：1000-3428(2021)06-0032-06 文献标志码：A 中图分类号：TP391

可验证的云存储医疗加密数据统计分析方案

张晓均[1,2]，张经伟[1]，黄 超[1]，唐 伟[1]
（1.西南石油大学 计算机科学学院，成都 610500；2.西南石油大学 网络空间安全研究中心，成都 610500）

摘 要：为满足当前云存储医疗数据对敏感性、完整性以及统计分析可用性的需求，提出一种可验证的医疗加密数据统计分析方案。采用同态加密技术实现密文数据聚合并提高医疗数据的机密性，通过同态签名算法确保外包医疗加密数据的完整性。用户上传经过同态加密和签名的医疗数据到云服务器，云服务器在收到医疗数据分析中心的外包数据聚合请求后对密文医疗数据以及签名值进行聚合运算，并将相应结果返回给医疗数据分析中心，医疗数据分析中心验证云服务器外包同态加密数据聚合的完整性。在此基础上，医疗数据分析中心仅需使用私钥解密就能获得所有用户正确的原始医疗数据聚合结果，并据此进行统计分析。实验结果表明，该方案在医疗隐私大数据分析领域相对SPPDA等方案具有效率优势，医疗数据分析中心在验证数据完整性和分析聚合数据时计算开销保持恒定，与用户数量无关。

关键词：云存储；医疗数据；加密聚合；同态加密；同态签名；完整性验证

开放科学（资源服务）标志码（OSID）：

中文引用格式：张晓均，张经伟，黄超，等.可验证的云存储医疗加密数据统计分析方案[J].计算机工程，2021，47(6)：32-37，43.

英文引用格式：ZHANG Xiaojun，ZHANG Jingwei，HUANG Chao，et al.Verifiable statistical analysis scheme for encrypted medical data in cloud storage[J].Computer Engineering，2021，47(6)：32-37，43.

Verifiable Statistical Analysis Scheme for Encrypted Medical Data in Cloud Storage

ZHANG Xiaojun[1,2]，ZHANG Jingwei[1]，HUANG Chao[1]，TANG Wei[1]
（1.School of Computer Science，Southwest Petroleum University，Chengdu 610500，China；
2.Research Center of Cyberspace Security，Southwest Petroleum University，Chengdu 610500，China）

【Abstract】In order to meet the requirements of medical data in cloud for sensitivity，integrity and statistical analysis applicability，this paper proposes a verifiable statistical analysis scheme for encrypted medical data.The scheme employs the homomorphic encryption technique to achieve medical data confidentiality and encrypted data aggregation. In addition，the homomorphic signature algorithm is used to ensure the integrity of outsourced medical data.The scheme enables users to upload the encrypted medical data and the corresponding signatures to the cloud server for storage.Once receiving a request for outsourced data aggregation from a medical Data Analysis Center（DAC），the cloud server aggregates those encrypted data and the corresponding signatures，and returns the results to DAC.The DAC could verify the integrity of the encrypted data aggregated by the cloud server.By the private key for decryption，DAC could directly obtain correct results of the aggregated original medical data of all users，and further perform statistical analysis.The experimental results show that the calculation cost of DAC in this scheme is constant and independent of the number of users in data integrity verification and aggregated data analysis，and the proposed scheme is more efficient than SPPDA and other schemes in massive private medical data analysis.

【Key words】cloud storage；medical data；encrypted aggregation；homomorphic encryption；homomorphic signature；integrity verification

DOI：10. 19678/j. issn. 1000-3428. 0058999

基金项目：国家自然科学基金（61902327，61872060）；金融数学福建省高校重点实验室（莆田学院）开放课题（JR201903）；西南石油大学青年科技创新团队项目（2019CXTD05）。

作者简介：张晓均（1985—），男，副教授、博士，主研方向为密码学、信息安全；张经伟、黄 超、唐 伟，硕士研究生。

收稿日期：2020-07-20 **修回日期：**2020-10-24 **E-mail：**zhangxjdzkd2012@163.com

0 概述

移动互联、物联网、云计算和大数据等新兴信息技术与信息感知方式的快速发展,改变了传统的医疗与健康服务模式[1-3]。近年来,随着移动医疗等新技术的应用,电子健康档案、临床检测数据、可穿戴传感器感知的个人健康状态记录等医疗数据都呈现爆炸式增长[4-6]。医疗大数据处理技术在临床决策支持系统、远程病人监控数据以及对病人健康档案等精准分析方面发挥着重要作用,已成为提高诊疗效率、减少可避免的人为误差与缓解医疗资源分布不均问题的有效途径。同时,云计算技术因其高效的计算能力和强大的存储空间,可以有效集成在无线医疗网络环境中,减缓医疗数据剧增所带来的存储和处理压力[7-9]。

尽管云计算技术在管理健康医疗大数据方面呈现明显优势,但是外包云存储医疗数据容易遭受各种安全性威胁[10-12],其中,受影响最大的是数据的机密性[13]。事实上,用户健康医疗数据的敏感性导致其往往以密文形式存储在云服务器,这将失去部分甚至大部分数据的可用性。因此,如何在数据隐私得到有效保护的情况下对云存储外包医疗数据进行快速医学统计分析,具有重要的研究价值和实际意义。

加密数据聚合[14]可有效促进具有隐私保护的医疗数据的统计分析。密码学中具有加法同态特性的加密算法首先被集成到聚合方案中[15],然后通过云服务器对大量的医疗数据密文进行同态聚合,并将聚合后的密文发送给医疗数据分析中心(DAC),同时有效降低通信带宽,最后在医疗数据分析中心端进行聚合数据解密,进一步进行具有隐私保护的大数据的统计分析。此外,数据完整性在医疗云存储环境应用中也极为关键[16-17],原因是恶意敌手为了某种利益可能在用户和云服务器的传输信道中截取数据并执行替换或篡改攻击。由于云服务器同时要处理来自不同用户的海量医疗数据,在进行加密聚合的过程中,可能会因操作失误导致错误地聚合原始外包密文数据,这样最终返回的聚合数值并非真实结果。因此,可验证的加密数据聚合方案是医疗数据分析中心进行大数据深度准确统计分析的有效保证。

近年来,已出现各种加密聚合方案,但这些方案用在医疗数据统计分析领域还相对较少[18-19],而且很多方案缺乏可验证功能[20-21]。一些可以部署在无线医疗网络且支持可验证功能的加密聚合方案[22-23],由于在设计过程中需要的双线性对计算开销与原始用户数量呈线性增长趋势,因此方案效率不高。

本文提出一种可验证的外包云存储医疗加密数据统计分析方案,该方案采用改进的BGN同态加密算法使云服务器可以对密文进行聚合运算,从而减轻医疗数据分析中心的计算压力。同时,设计一种基于椭圆曲线的同态数字签名算法,使医疗数据分析中心在使用云端聚合的数据时,只需执行恒定的运算量即可高效地验证加密医疗聚合数据的完整性。

1 预备知识

1.1 双线性对

基于椭圆曲线的双线性对定义如下:

定义1 给定一个双线性对映射 $\tilde{e}: G_1 \times G_1 \to G_2$,其中,加法循环群 G_1 与乘法循环群 G_2 有共同的阶 q_1,且 V 是 G_1 的生成元。基于椭圆曲线的双线性对满足以下性质:

1)双线性。对任意2个群上的元素 $P、V \in G_1$ 和任意的 $a、b \in Z_q^*$,双线性对运算满足 $\tilde{e}(aP, bV) = \tilde{e}(P, V)^{ab}$。

2)非退化性。对于加法循环群上存在的2个元素 $P、V \in G_1$,$\tilde{e}(P, V) \neq 1 \in G_2$,此处1是 G_2 的单位元。

3)可计算性。对于加法循环群上任意的2个元素 $P、V \in G_1$,都能找到一个有效的算法计算 $\tilde{e}(P, V)$。

1.2 BGN公钥加密系统

BGN公钥加密系统[24]主要包括密钥生成、加密和解密3个算法。

1)密钥生成。给定安全参数 κ,合数阶双线性对产生器输出 (n, g, G, G_T, e),其中,$n = pq$,$p、q$ 是长度为 κ bit 的大素数,$G、G_T$ 是2个阶为 n 的循环群,g 是 n 阶循环群 G 的生成元,$e: G \times G \to G_T$ 是一个非退化的双线性对映射。设置 $u = g^p$,则 u 是 G 的 q 阶循环子群的生成元。公钥是 $pk = (n, G, G_T, e, g, u)$,私钥是 $sk = q$。

2)加密。假设明文空间是一个整数集 $\{1, 2, \cdots, T\}$,其中,$T < p$。选取随机数 $r \in Z_n$,明文 m 的密文 $c = \mathrm{Enc}(m, r) = g^m u^r \in G$。

3)解密。给定密文 $c = g^m u^r$,利用私钥 q 计算 $c^q = (g^m u^r)^q = (g^q)^m$,然后计算以 g^q 为底的 c^q 的离散对数,即可恢复 m。由于 $0 \leqslant m \leqslant T$,根据文献[24]中Pollard的lambda解密方法,BGN公钥加密算法可在 $O(\sqrt{T})$ 时间复杂度内解密出明文。

2 方案设计

2.1 系统模型

可验证的外包云存储医疗加密数据聚合方案的系统模型,包含用户、云服务器、医疗数据分析中心和可信中心(TA)4个通信实体,如图1所示。

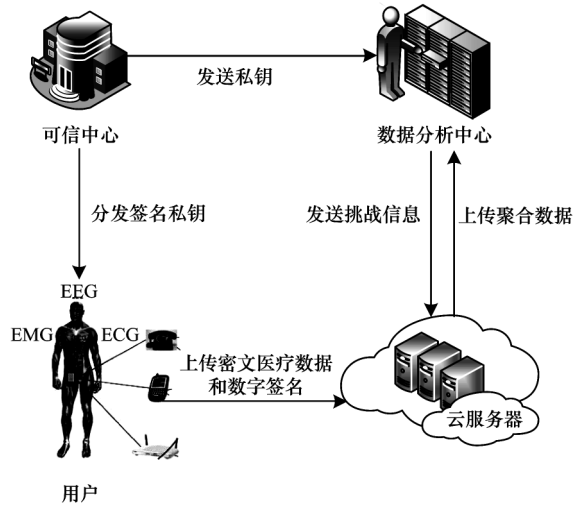

图1 加密数据聚合方案的系统模型

Fig.1 System model of encrypted data aggregation scheme

系统模型中的4个通信实体具体如下：

1)用户。通过可穿戴设备收集健康医疗数据，使用移动终端计算设备对医疗数据进行加密，产生密文对应的数字签名，最后将所有密文及对应的数字签名集合上传到远程云服务器。

2)云服务器。拥有巨大的计算和存储能力，在本模型中主要用于存储用户上传的加密数据及数字签名集合。一旦接收到医疗数据分析中心的挑战请求，云服务器会对挑战位置的加密数据及数字签名进行同态聚合运算，并返回最终结果到医疗数据分析中心。

3)医疗数据分析中心。当接收到来自云服务器返回的聚合结果时，医疗数据分析中心首先进行加密数据完整性验证，然后利用私钥来解密以获得不同用户医疗数据的聚合值，最后对用户的医疗数据进行隐私保护统计分析。

4)可信中心。负责设置并发布系统的公开密码参数，系统初始化阶段通过安全信道为各通信实体发送私钥。

本文提出的可验证外包云存储医疗加密数据聚合方案，重点解决云存储医疗数据的机密性、完整性以及数据统计分析可用性问题。因此，本系统中引入的可信中心实际上需要使用相关身份认证技术[25-27]，对每个通信实体进行身份验证后才能进入医疗云存储系统并为其颁发对应的公私钥。

2.2 具体步骤

可验证外包云存储医疗加密数据聚合方案具体包括4个阶段：系统初始化，医疗数据加密和签名上传，加密医疗数据同态聚合，验证和聚合数据解密。

1)系统初始化。可信中心TA生成用于同态加密、同态数字签名和验证的系统公共参数。同时，TA将秘密参数发送给医疗数据分析中心以及对应的用户。系统初始化阶段具体步骤如下：

(1) TA选取长度相等的大素数p_1和p_2，满足$n=p_1p_2$，设置双线性对映射$e: G_a\times G_a\rightarrow G_b$，其中，$G_a$为$n$阶乘法循环群，$g$为$G_a$的生成元，选取$G_a$的$p_1$阶子群的生成元$x=g^{p_2}$。

(2) TA选取有限域F_p(p是大素数)上的椭圆曲线E，并基于该椭圆曲线设置另外一个双线性对映射$\tilde{e}$: $G_1\times G_1\rightarrow G_2$，$V$是基于椭圆曲线$E$的$q$阶加法循环群$G_1$的生成元。TA设置需要外包到云服务器的具有某一类型的医疗数据的用户数量为N，对于第i个用户，TA为其生成私钥$z_i\in Z_q$，并计算公钥$U_i=z_iV$。TA设置2个抗碰撞的哈希函数H_1:$\{0,1\}^*\rightarrow G_1$，H_2:$\{0,1\}^*\rightarrow Z_q^*$。

最后，TA公开如下的系统公共参数：

$$\text{pub}=(G_a, G_b, e, g, x, n, G_1, G_2, \tilde{e}, V, \{U_i\}_{1\leqslant i\leqslant N}, H_1, H_2)$$

TA通过安全信道将私钥p_1发送给医疗数据分析中心，将私钥z_i发送给对应的用户i。

2)医疗数据加密和签名上传。用户i利用医疗数据分析中心的公钥生成医疗数据m_i的密文，同时使用私钥对密文数据产生对应的数字签名σ_i，最后将密文和对应的数字签名数据上传到云服务器。医疗数据加密和签名上传阶段具体步骤如下：

(1) 对于明文数据m_i，其最大值T小于p_2，随机选取$s_i\in Z_n$，计算密文$c_i=\text{Enc}(m_i,s_i)=g^{m_i}x^{s_i}\in G_a$。

(2) 计算数字签名$\sigma_i=(z_i+H_2(c_i))H_1(\text{type})$，其中，type是医疗数据的类型。

最后，每一个用户i将自己生成的签名数据和密文数据$\{\sigma_i,c_i\}$一起上传到远程云服务器。

3)加密医疗数据同态聚合：当医疗数据分析中心需要分析某一类型的敏感医疗数据时，使用伪随机数发生器生成含l个伪随机数的伪随机序列$\{t_1,t_2,\cdots,t_{l-2},\alpha,\beta\}$，将医疗数据类型type和伪随机序列一起作为挑战信息chal发送给云服务器。然后，云服务器根据type类型医疗数据上N个用户的外包密文数据和这些数据对应的数字签名，分别进行聚合。加密医疗数据同态聚合阶段具体步骤如下：

(1) 云服务器对N个加密数据进行同态聚合：

$$\text{SC}=\prod_{i=1}^{N}C_i=\prod_{i=1}^{N}\text{Enc}(m_i,s_i)=\prod_{i=1}^{N}g^{m_i}x^{s_i}=g^{\sum_{i=1}^{N}m_i}\cdot x^{\sum_{i=1}^{N}s_i}$$

(2) 根据双线性对的运算性质和同态加密性质，对每个密文$c_i=\text{Enc}(m_i,s_i)$进行如下聚合：

$$\begin{aligned}\text{QSC}&=\prod_{i=1}^{N}e(c_i,c_i)=\\&\prod_{i=1}^{N}e(\text{Enc}(m_i,s_i),\text{Enc}(m_i,s_i))=\\&\prod_{i=1}^{N}e(g^{m_i}x^{s_i},g^{m_i}x^{s_i})=\\&\prod_{i=1}^{N}e(g,g)^{m_i^2}\cdot e(g,x)^{2m_is_i+p_2s_i^2}\end{aligned}$$

(3)基于以上2个聚合数据值SC、QSC和挑战信息，云服务器利用哈希函数H_2产生新的随机数$t_{l-1}=H_2(\text{SC}\|\alpha)$和$t_l=H_2(\text{QSC}\|\beta)$，云服务器进一步基于挑战信息中的伪随机序列$\{t_1,t_2,\cdots,t_{l-2},t_{l-1},t_l\}$对$N$个数字签名数据$\sigma_1,\sigma_2,\cdots,\sigma_N$进行聚合，$\sigma=\sum_{i=1}^{N}(t_jH_1(\text{type})+\sigma_i)$，其

中，$i\in\{1,2,\cdots,N\}$，$j\in\{1,2,\cdots,l\}$，$l<N$，$j=(i-1)\bmod l+1$。云服务器计算 $c=\sum_{i=1}^{N}H_2(c_i)$，将对应公钥 $\{U_1,U_2,\cdots,U_N\}$ 进行聚合，$U=\sum_{i=1}^{N}U_i$。

最后，云服务器将所有聚合数据 Agg={σ, c, U, SC, QSC}发送给医疗数据分析中心。

4）验证和聚合数据解密：当接收到云服务器发送的聚合数据后，医疗数据分析中心执行数据完整性验证，并对聚合密文SC和QSC进行解密。验证和聚合数据解密阶段具体步骤如下：

（1）计算 $t_{l-1}=H_2(\mathrm{SC}\|\alpha)$、$t_l=H_2(\mathrm{QSC}\|\beta)$ 以及 $t=\sum_{i=1}^{N}t_j$，其中，$j=(i-1)\bmod l+1$，验证如下方程是否成立：

$$\tilde{e}(\sigma,V)=\tilde{e}((c+t)H_1(\mathrm{type}),V)\cdot\tilde{e}(H_1(\mathrm{type}),U)$$

（2）一旦验证上述方程成立，医疗数据分析中心确信外包云存储密文数据未被云服务器错误聚合或者被外部敌手恶意替换、篡改。根据文献[24]中Pollard的lambda解密方法，医疗数据分析中心利用私钥 p_1 进行条件性穷举暴力破解，在时间复杂度为 $O(\sqrt{T})$ 的情况下可有效求解离散对数 $\log_{g^{p_1}}\mathrm{SC}^{p_1}$，进而恢复该类医疗数据的统计和 $\sum_{i=1}^{N}m_i$。同样，医疗数据分析中心可有效求解离散对数 $\log_{e(g,g)^{p_1}}\mathrm{QSC}^{p_1}$，恢复医疗数据的平方和 $\sum_{i=1}^{N}m_i^2$。

2.3　加密数据聚合的正确性及统计分析

医疗数据分析中心通过检验完整性验证方程是否成立，以判断云服务器是否按照正确步骤进行同态加密聚合。完整性验证方程推导如下：

$$\tilde{e}(\sigma,V)=\tilde{e}\left(\sum_{i=1}^{n}(t_jH_1(\mathrm{type})+\sigma_i),V\right)=$$
$$\tilde{e}\left(\sum_{i=1}^{n}(t_j+z_i+H_2(c_i))H_1(\mathrm{type}),V\right)=$$
$$\tilde{e}\left(\sum_{i=1}^{n}(t_j+H_2(c_i))H_1(\mathrm{type}),V\right)\cdot$$
$$\tilde{e}\left(\sum_{i=1}^{n}(z_iH_1(\mathrm{type}),V\right)=$$
$$\tilde{e}\left(\left(\sum_{i=1}^{n}H_2(c_i)+t\right)H_1(\mathrm{type}),V\right)\cdot$$
$$\tilde{e}\left(\sum_{i=1}^{n}z_iH_1(\mathrm{type}),V\right)=$$
$$\tilde{e}\left((c+t)H_1(\mathrm{type}),V)\cdot\tilde{e}(H_1(\mathrm{type}),\sum_{i=1}^{n}z_iV\right)=$$
$$\tilde{e}((c+t)H_1(\mathrm{type}),V)\cdot\tilde{e}(H_1(\mathrm{type}),U)=$$
$$\tilde{e}\left(\left(\sum_{i=1}^{n}H_2(c_i)+\sum_{i=1}^{n}t_j\right)H_1(\mathrm{type}),V\right)\cdot$$
$$\tilde{e}\left(\sum_{i=1}^{n}z_iH_1(\mathrm{type}),V\right)$$

当医疗数据分析中心得到正确的type类型医疗数据的统计和 $\sum_{i=1}^{N}m_i$ 以及平方和 $\sum_{i=1}^{N}m_i^2$ 时，其可计算出该类医疗数据的平均值和方差统计数据，分别如下：

$$\bar{m}=\frac{1}{N}\cdot\sum_{i=1}^{N}m_i$$

$$\mathrm{var}(m_i)=\frac{1}{N}\cdot\sum_{i=1}^{N}m_i^2-\left(\frac{1}{N}\cdot\sum_{i=1}^{N}m_i\right)^2$$

最后，医疗数据分析中心根据上述统计数据，可在保护用户医疗数据隐私的情况下进行大数据处理和深度分析。图2所示为详细的数据聚合与统计分析流程。

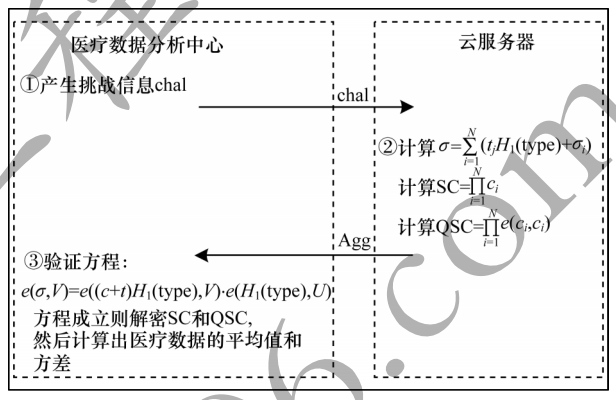

图2　数据聚合与统计分析流程

Fig.2　Procedure of data aggregation and statistical analysis

3　安全性证明

定理1　可验证的外包云存储医疗加密数据聚合方案可确保用户外包云存储医疗数据的机密性。

证明　每一个移动终端用户 i 产生医疗数据 m_i 的密文 $c_i=g^{m_i}x^{s_i}$，发送到云服务器，c_i 本质上是改进的BGN同态加密系统的密文，由于此密码系统满足选择明文安全的语义安全性，即便敌手在用户和云服务器的公开信道截获到相关密文，也不能恢复用户的原始医疗数据。此外，一旦接收到来自 N 个用户的所有加密医疗数据 c_i，$i=1,2,\cdots,N$，云服务器则对 N 个加密数据进行同态聚合：

$$\mathrm{SC}=\prod_{i=1}^{N}c_i=\prod_{i=1}^{N}g^{m_i}x^{s_i}=g^{\sum_{i=1}^{N}m_i}\cdot x^{\sum_{i=1}^{N}s_i}$$

$$\mathrm{QSC}=\prod_{i=1}^{N}e(c_i,c_i)=\prod_{i=1}^{N}e(g^{m_i}x^{s_i},g^{m_i}x^{s_i})=$$
$$\prod_{i=1}^{N}e(g,g)^{m_i^2}\cdot e(g,x)^{2m_is_i+p_2s_i^2}$$

经过上述过程，这 N 个用户的加密医疗数据被云服务器聚合为SC和QSC 2个密文，两者本质上也分别是 $\sum_{i=1}^{N}m_i$ 和 $\sum_{i=1}^{N}m_i^2$ 的改进BGN加密系统的密文。同样，根据BGN加密系统选择明文的语义安全性，即便敌手在云服务器和医疗数据分析中心的公开信道截获到这2个聚合密文，也不能恢复用户原始医

疗数据的统计和与平方和。

定理2 可验证的外包云存储医疗加密数据聚合方案可确保云存储加密医疗数据聚合的可验证性。

证明 在医疗数据加密和签名上传阶段，加密医疗数据的数字签名 $\sigma_i=(z_i+H_2(c_i))H_1(\text{type})$ 是用户利用自己的私钥 z_i 产生的，任意敌手如果能在多项式时间内伪造一个新的数字签名 σ_i^*，其必然在多项式时间内可以求解基于椭圆曲线的离散对数困难问题。因此，单个密文的数字签名伪造是不可行的。同样，敌手在多项式时间内伪造聚合的加密医疗数据数字签名 $\sigma=\sum_{i=1}^{N}(t_jH_1(\text{type})+\sigma_i)(j=(i-1)\bmod l+1)$ 也是不可行的。此外，由于 $c_i=g^{m_i}x^{s_i}$ 是用户利用医疗数据分析中心的公钥产生的，敌手自己可能产生一个替换的密文 c_i^*，因此云服务器在产生聚合密文信息 $\text{SC}=\prod_{i=1}^{N}c_i$ 和 $\text{QSC}=\prod_{i=1}^{N}e(c_i,c_i)$ 之后，在返回给医疗数据分析中心的过程中，2个信息可能被替换为 SC^* 和 QSC^*。于是，篡改的聚合信息 $\text{Agg}^*=\{\sigma,c,\text{SC}^*,\text{QSC}^*\}$ 要通过医疗数据分析中心的验证，其必须满足如下方程：

$$\tilde{e}(\sigma,V)=\tilde{e}\left(\left(c+\sum_{i=1}^{l-2}t_i+t_{l-1}^*+t_l^*\right)H_1(\text{type}),V\right)\cdot\tilde{e}(H_1(\text{type}),U)$$

其中，$t_{l-1}^*=H_2(\text{SC}^*\|\alpha)$，$t_l^*=H_2(\text{QSC}^*\|\beta)$。而正确的聚合信息 $\text{Agg}=\{\sigma,c,\text{SC},\text{QSC}\}$ 应满足如下的验证方程：

$$\tilde{e}(\sigma,V)=\tilde{e}((c+t)H_1(\text{type}),V)\cdot\tilde{e}(H_1(\text{type}),U)$$

根据以上2个验证方程得知：

$$\left(c+\sum_{i=1}^{l-2}t_i+t_{l-1}^*+t_l^*\right)H_1(\text{type})=(c+t)H_1(\text{type})$$

设置 $W=(t_{l-1}+t_l)H_1(\text{type})$，可以求解 $H_1(\text{type})$ 和 W 之间的离散对数 $t_{l-1}^*+t_l^*$，这与基于椭圆曲线的离散对数困难问题假设是矛盾的。因此，根据以上安全性分析得知本文方案可确保云存储加密医疗数据聚合的可验证性，即医疗数据分析中心可以验证云服务器加密聚合过程的正确性以及聚合密文的完整性。

4 性能分析

将本文方案与文献[22-23]中的2种可验证加密聚合方案进行性能分析与对比，所有方案都运行在处理器为Inter®Core™i5-2320 3.00 GHz和内存8.00 GB的主机上，操作系统为Windows10。所有方案均通过C语言以及版本号为5.6.2的密码算法基础函数库MIRACL来实现，使用的椭圆曲线密码机制是MNT曲线，嵌入阶是6。定义 T_{pa} 表示双线性对运行时间，T_{mu} 表示普通乘法运行时间，T_{Mu} 表示椭圆曲线上加法循环群中倍点计算运行时间，T_{Ad} 表示椭圆曲线上加法循环群中加法的运行时间，T_{ex} 表示普通模指数计算运行时间，T_{Ex} 表示双线性对映射中的模指数计算运行时间，T_{Ha} 表示映射到椭圆曲线上加法循环群的哈希函数运行时间，T_{ha} 表示普通哈希函数的运行时间。

在本文系统模型中，各种密文同态聚合和数字签名同态聚合运算都外包给具有强大计算能力的远程云服务器，同时，各个对比方案都是通过聚合计算来降低通信开销的。因此，本节通过比较各方案在医疗数据分析中心端的数据处理中所需的计算开销，以验证本文方案在计算性能上的优势。具体地，根据分析得知，文献[22]中的SPPDA方案在验证和聚合数据解密过程中，数据中心需要执行 $N+1$ 个双线性对运算和 N 个普通的哈希函数，才能通过完整性验证方程，同时，数据中心需要执行1个双线性对运算、1个普通模指数计算和1个普通乘法，才能实现完整聚合密文的解密。因此，SPPDA方案中数据中心的总计算开销为 $(N+2)T_{\text{pa}}+T_{\text{ex}}+NT_{\text{ha}}+T_{\text{mu}}$。文献[23]方案在验证和聚合数据解密过程中，数据中心需要执行 $N+1$ 个双线性对运算、N 个普通的哈希函数以及 $N-1$ 个普通乘法，才能通过完整性验证方程，同时，数据中心需要执行1个普通模指数计算实现聚合密文的解密。因此，文献[23]方案数据中心的总计算开销为 $(N+1)T_{\text{pa}}+NT_{\text{ha}}+(N-1)T_{\text{mu}}+T_{\text{ex}}$。在本文方案中，医疗数据分析中心需要执行3个双线性对运算、1个椭圆曲线上加法循环群中的倍点计算、1个映射到椭圆曲线上加法循环群的哈希函数以及2个普通的哈希函数，才能通过完整性验证方程，同时，医疗数据分析中心需要执行1个普通模指数计算和1个双线性对映射中的模指数计算，才能实现聚合密文的解密。因此，本文方案中医疗数据分析中心总的计算开销为 $3T_{\text{pa}}+T_{\text{Mu}}+T_{\text{Ha}}+2T_{\text{ha}}+T_{\text{ex}}+T_{\text{Ex}}$。3种方案的验证和聚合数据解密过程具体计算开销如表1所示。

表1 3种方案的数据分析中心计算开销对比

Table 1 Comparison of computing cost of data analysis center of three schemes

方案	验证过程	解密过程	总计算开销
文献[22]方案	$(N+1)T_{\text{pa}}+NT_{\text{ha}}$	$T_{\text{pa}}+T_{\text{ex}}+T_{\text{mu}}$	$(N+2)T_{\text{pa}}+T_{\text{ex}}+NT_{\text{ha}}+T_{\text{mu}}$
文献[23]方案	$(N+1)T_{\text{pa}}+NT_{\text{ha}}+(N-1)T_{\text{mu}}$	T_{ex}	$(N+1)T_{\text{pa}}+NT_{\text{ha}}+(N-1)T_{\text{mu}}+T_{\text{ex}}$
本文方案	$3T_{\text{pa}}+T_{\text{Mu}}+T_{\text{Ha}}+2T_{\text{ha}}$	$T_{\text{ex}}+T_{\text{Ex}}$	$3T_{\text{pa}}+T_{\text{Mu}}+T_{\text{Ha}}+2T_{\text{ha}}+T_{\text{ex}}+T_{\text{Ex}}$

从图3可以看出，本文方案医疗数据分析中心在计算效率方面具有明显优势，特别地，随着移动终端用户数目N的增加，文献[22-23]方案的计算开销均呈线性增长，而本文方案医疗数据分析中心的计算开销保持恒定轻量级常量，并未随着移动终端用户数目N的增加而增加。

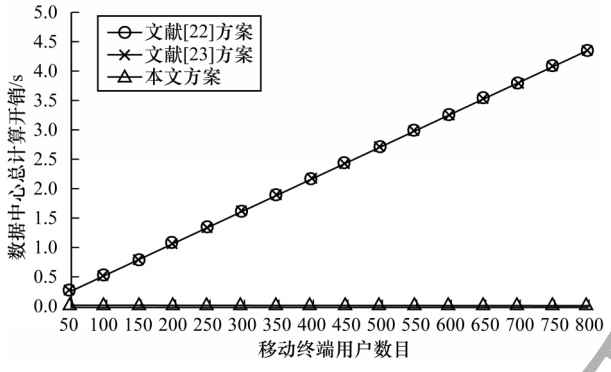

图3　3种方案的数据分析中心计算效率比较

Fig.3　Comparison of computing efficiency of data analysis center of three schemes

5　结束语

医疗数据作为用户的隐私数据，在医生对患者的病情判断中发挥重要作用，在统计分析医疗数据时必须保证其机密性和完整性。本文提出一种可验证的云存储医疗加密数据统计分析方案，基于改进的BGN同态加密算法，在保障医疗数据机密性的同时，将加密数据聚合操作外包给云服务器，以缓解医疗数据分析中心的计算压力。同时，设计一种基于椭圆曲线的数字签名算法，使医疗数据分析中心可以快速验证云服务器所聚合医疗数据的真实性。性能分析结果表明，该方案在医疗数据分析中心只需轻量级恒定计算开销的情况下就能判断数据在传输和存储期间是否遭受篡改或替换，并据此对医疗数据进行均值和方差等统计分析。但本文研究尚未考虑用户可能不愿意上传敏感医疗数据、用户数据传输中途被中断或者其他恶意攻击行为导致数据传输失败等实际情况，因此，下一步将增加容错机制并采用门限秘密共享技术，使得本文方案在有效传输数据达到门限值的情况下即可顺利完成加密数据聚合。

参考文献

[1] CAVALLARI R, MARTELLI F, ROSINI R, et al. A survey on wireless body area networks: technologies and design challenges[J]. IEEE Communications Surveys and Tutorials, 2014, 16(3): 1635-1657.

[2] DEY N, ASHOUR S A, SHI F Q, et al. Medical cyber-physical systems: a survey[J]. Journal of Medical Systems, 2018, 42(74): 1-13.

[3] ARTHUR G, YOUAKIM B, BERTR M, et al. Internet of medical things: a review of recent contributions dealing with cyber-physical systems in medicine[J]. IEEE Internet of Things Journal, 2018, 5(5): 3810-3822.

[4] LU Rongxing, ZHU Hui, LIU Ximeng, et al. Toward efficient and privacy-preserving computing in big data era[J]. IEEE Network, 2014, 28(4): 46-50.

[5] WU Xindong, ZHU Xingquan, WU Gongqing, et al. Data mining with big data[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(1): 97-107.

[6] ZHANG Y, QIU M K, TSAI C W, et al. Health-CPS: healthcare cyber-physical system assisted by cloud and big data[J]. IEEE Systems Journal, 2017, 11(1): 88-95.

[7] WAN J F, ZOU C F, ULLAH S, et al. Cloud-enabled wireless body area networks for pervasive healthcare[J]. IEEE Network, 2013, 27(5): 56-61.

[8] ULLAH S, VASILAKOS A, SUZUKI J, et al. Cloud-assisted wireless body area networks[J]. Information Sciences, 2014, 28(4): 81-83.

[9] ZHANG Yuan, XU Chunxiang, LI Hongwei, et al. HealthDep: an efficient and secure deduplication scheme for cloud-assisted eHealth systems[J]. IEEE Transactions on Industrial Informatics, 2018, 14(9): 4101-4112.

[10] WANG Huaqun, WU Qianhong, QIN Bobo, et al. FRR: fair remote retrieval of put sourced private medical records in electronic health networks[J]. Journal of Biomedical Informatics, 2014, 50(8): 226-233.

[11] ZHOU Jun, CAO Zhenfu, DONG Xiaolei, et al. Security and privacy in cloud-assisted wireless wearable communications: challenges, solutions, and future directions[J]. IEEE Wireless Communications, 2015, 22(2): 136-144.

[12] NHLABATSI A M, HONG J B, KIM D S, et al. Threat-specific security risk evaluation in the cloud[EB/OL]. [2020-06-20]. https://ieeexplore.ieee.org/document/8543671.

[13] ZHANG Y H, ZHENG D, DENG R H. Security and privacy in smart health: efficient policy-hiding attribute-based access control[J]. IEEE Internet of Things Journal, 2018, 5(3): 2130-2145.

[14] LU Rongxing, LIANG Xiaohui, LI Xu, et al. EPPA: an efficient and privacy-preserving aggregation scheme for secure smart grid communications[J]. IEEE Transactions on Parallel & Distributed Systems, 2012, 23(9): 1621-1632.

[15] CHEN Zhiwei, DU Min, YANG Yatao, et al. Homomorphic cloud computing scheme based on RSA and Paillier[J]. Computer Engineering, 2013, 39(7): 35-39. (in Chinese)
陈志伟，杜敏，杨亚涛，等. 基于RSA和Paillier的同态云计算方案[J]. 计算机工程，2013，39(7)：35-39.

[16] HE D B, ZEADALLY S, WU L B. Certificateless public auditing scheme for cloud-assisted wireless body area networks[J]. IEEE Systems Journal, 2018, 12(1): 64-73.

[17] ZHANG Xiaojun, ZHAO Jie, XU Chunxiang, et al. CIPPPA: conditional identity privacy-preserving public auditing for cloud-based WBANs against malicious auditors[J]. IEEE Transactions on Cloud Computing, 2019, 15(6): 15-39.

[18] GUI Yihong. Research on HEDSA data aggregation of wireless sensor networks[J]. Computer Engineering, 2011, 37(7): 160-162. (in Chinese)
归奕红. 无线传感器网络HEDSA数据聚合研究[J]. 计算机工程，2011，37(7)：160-162.

（下转第43页）

（上接第37页）

[19] ZHANG K, LIANG X H, BAURA M, et al. PHDA: a priority based health data aggregation with privacy preservation for cloud assisted WBANs[J]. Information Sciences, 2014, 284: 130-141.

[20] HAN Song, ZHAO Shuai, LI Qinghua, et al. PPM-HDA: privacy-preserving and multifunctional health data aggregation with fault tolerance for cloud assisted WBANs[J]. IEEE Transactions on Information Forensics & Security, 2016, 11(9): 1940-1955.

[21] LI R N, STURTIVANT C, YU J G, et al. A novel secure and efficient data aggregation scheme for IoT[J]. IEEE Internet of Things Journal, 2018, 6(2): 1551-1560.

[22] ARA A, AL-RODHAAN M, TIAN Y, et al. SPPDA scheme based on bilinear elgamal cryptosystem [J]. IEEE Access, 2017, 5: 12601-12617.

[23] LI Xiong, LIU Shanpeng, WU Fan, et al. Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications[J]. IEEE Internet of Things Journal, 2019, 6(3): 4755-4763.

[24] LU Rongxing. Privacy-enhancing aggregation techniques for smart grid communications[M]. Berlin, Germany: Springer, 2016.

[25] WANG Chenyu, WANG Ding, WANG Feifei, et al. Multi-factor user authentication scheme for multi-gateway wireless sensor networks[J]. Chinese Journal of Computers, 2020, 43(4): 683-700. (in Chinese)
王晨宇, 汪定, 王菲菲, 等. 面向多网关的无线传感器网络多因素认证协议[J]. 计算机学报, 2020, 43(4): 683-700.

[26] WANG Ding, WANG Ping, LEI Ming. Cryptanalysis and improvement of gateway-oriented password authenticated key exchange protocol based on RSA[J]. Acta Electronica Sinica, 2015, 43(1): 176-184. (in Chinese)
汪定, 王平, 雷鸣. 基于RSA的网关口令认证密钥交换协议的分析与改进[J]. 电子学报, 2015, 43(1): 176-184.

[27] WANG Ding, WANG Peng. Two birds with one stone: two-factor authentication with security beyond conventional bound[J]. IEEE Transactions on Dependable and Secure Computing, 2016, 15(4): 708-722.

编辑　吴云芳