基于区块链的分布式 EHR 细粒度可追溯方案

应作斌[1]，斯元平[2]，马建峰[2,3]，刘西蒙[4]

（1. 安徽大学计算机科学与技术学院，安徽 合肥 230601；2. 安徽大学物质科学与信息技术研究院，安徽 合肥 230601；
3. 西安电子科技大学网络与信息安全学院，陕西 西安 710071；4. 福州大学数学与计算机科学学院，福建 福州 350108）

摘 要：针对电子健康档案（EHR）在分布式系统中的密钥管理及用户身份追溯问题，提出了一种基于区块链的分布式 EHR 细粒度可追溯方案。结合变色龙哈希和零知识证明技术实现区块链上节点的注册与身份证明的生成，从而实现区块链上恶意用户的追溯。针对密钥管理的单点故障问题，设计了分布式密文策略的属性基加密方案实现安全细粒度的数据访问控制，设置多个解密机构区块链节点联合分发用户节点的属性私钥。安全性分析表明，基于区块链的可追溯分布式密钥生成属性基加密算法是随机预言机模型下自适应安全的，并通过实验证明了所提方案的可行性和实用性。

关键词：电子健康档案；区块链；追溯；密钥管理；细粒度访问控制

中图分类号：TP309

文献标识码：A

DOI: 10.11959/j.issn.1000−436x.2021033

Blockchain-based distributed EHR fine-grained traceability scheme

YING Zuobin[1], SI Yuanping[2], MA Jianfeng[2,3], LIU Ximeng[4]

1. School of Computer Science and Technology, Anhui University, Hefei 230601, China
2. Institutes of Physical and Information Technology, Anhui University, Hefei 230601, China
3. School of Network and Information Security, Xidian University, Xi'an 710071, China
4. School of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China

Abstract: Aiming at the key management of electronic health records (EHR) in a distributed system and user identity tracing issues, a distributed EHR fine-grained traceability scheme based on blockchain was proposed. Combining chameleon hash and zero-knowledge proof technology, the registration of nodes on the blockchain and the generation of identity certificates were realized, and the traceability of malicious users on the blockchain was realized. Besides, given the single point of failure problem of key management, the attribute-based encryption scheme of distributed ciphertext strategy was designed to achieve secure and fine-grained data access control, and multiple decryption agency blockchain nodes were set up to jointly distribute the attribute private keys of user nodes. The security analysis shows that the traceable distributed key generation attribute-based encryption algorithm based on the blockchain is adaptively secure under the random oracle model, and through experiments, the feasibility and practicability of the proposed scheme are shown.

Keywords: EHR, blockchain, tracking, key management, fine-grained access control

1 引言

目前，电子健康档案（EHR, electronic health record）已经成为提高医疗诊断效率的工具之一，其数据从可穿戴设备、智能传感器等来源收集。随着智慧医疗的逐步发展，EHR 数据量呈指数级增长。电子病历的数据量将以每年 48%的速度增长[1]。但是，EHR 数据共享面临存储安全性和隐私泄露的

收稿日期：2020−08−15；**修回日期**：2020−11−10

基金项目：安徽省教育厅重点基金资助项目（No.KJ2018A0031）；国家自然科学基金资助项目（No.62072109, No.U1804263, No.61702105）

Foundation Items: The Key Project of Anhui Provincial Department of Education (No.KJ2018A0031), The National Natural Science Foundation of China (No.62072109, No.U1804263, No.61702105)

问题。密文策略属性基加密（CP-ABE, ciphertext policy attribute based encryption）专门为一对多加密而设计，适合作为 EHR 的访问控制解决方案。面对海量的 EHR 数据，本地存储解决方案已不再适用，而集中式云存储方案存在数据泄露和单点故障问题，区块链技术成为 EHR 数据管理的有前途的解决方案。考虑到区块链的块大小有限，将加密的 EHR 文件存储于分布式存储系统，如星际文件系统（IPFS, inter planetary file system），而将 EHR 的密文及其 IPFS 中的下载地址等数据上链，可以有效节省区块链的存储空间。

基于区块链的 EHR 访问控制方案的研究大多侧重于隐私保护[2-3]，而实现隐私保护的同时带来了监管问题。区块链中强大的隐私保护滋生了许多安全问题，例如，所有勒索软件攻击中有 88%发生在医疗保健系统中[4]。文献[5]指出，2009—2017 年发生了约 1 138 起违规事件，影响了 1.64 亿患者的医疗数据。这些安全问题不仅对用户利益构成严重威胁，而且严重阻碍了区块链在 EHR 大规模共享中的开发和应用。跟踪用户身份是解决区块链监管问题的关键。

为了应对这些挑战，本文提出了一种基于区块链的分布式 EHR 细粒度可追溯方案，能够实现 EHR 在分布式存储中的隐私保护和细粒度的访问控制，并完成用户的追溯。此外，为了消除分布式存储中密钥管理的单点故障（SPoF, single point of failure）问题，本文设计了一种分布式 CP-ABE 方案。本文主要的贡献总结如下。

1) 提出了一种基于区块链的可追溯分布式 EHR 细粒度访问控制方案，解决分布式存储中 EHR 数据的密钥管理和恶意用户的追溯问题。

2) 融合变色龙哈希和零知识证明技术完成链上节点的注册验证和生成身份证明，通过节点的公开信息可以追溯节点的身份证明，获得注册节点的真实 ID，从而实现用户的追溯。

3) 结合分布式密钥生成（DKG, distributed key generation）协议，设定多个解密机构节点作为授权中心，设计分布式 CP-ABE 方案。安全性证明表明，基于区块链的可追溯分布式属性基加密算法在随机预言机模型下是自适应安全的。

2 相关工作

2008 年，中本聪在比特币中使用了哈希链和工作量证明[6]。区块链[7]是按时间排序的数据区块，包括完整和有效的交易记录列表。近年来，区块链技术已应用于金融服务[8]、医疗保健[2, 9]、物联网[10-11]和车联网[2, 12]等领域。将区块链应用于医疗保健领域的趋势正在增加[13-16]。文献[13]提出了使用区块链技术的去中心化电子病历管理系统——MedRec，它是一种模块化设计，可管理参与者之间的权限和数据共享。与 MedRec 类似，文献[14]提出了一个基于以太坊的区块链 Ancile，它利用智能合约来增强访问控制和数据混淆。文献[15]构建了云环境中电子病历的基于区块链的数据共享框架。文献[16]提出通过区块链进行高效、安全的医疗数据共享，利用基于混合区块链的架构保护电子病历。

在基于区块链的存储 EHR 数据的应用中，研究者提出加密方案来增强 EHR 的安全性和有效性。文献[17]采用离散波长变换和遗传算法来增强安全性并优化系统性能。类似地，文献[18]通过在密钥的共享中采用密文策略的属性基加密，并结合区块链技术，实现了云存储中的细粒度访问控制，并支持验证搜索。文献[19]将基于分层身份的加密系统（HIBE, hierarchical identity based encryption）和 CP-ABE 组合在一起，以实现在云服务器上有效的数据加密共享。文献[20]提出了一种用于大数据访问的有效可撤销 CP-ABE 方案，使用基于代理的更新在云中进行控制。Lewko 和 Waters[21]提出在没有中央授权的情况下提供分布式 CP-ABE 方案。Hu 等[22]和 Li 等[23]提出了具有隐私保护和基于双方属性的密钥协议的多权限 CP-ABE 方案。

但是，上述应用均集中在增强安全性和隐私保护方面，密钥生成的单点故障问题依然影响安全，且对于区块链的可追溯功能的研究是有限的。EHR 数据具有高度敏感性，在此类数据的共享中，一旦发生非法访问，将会造成极大的不良影响。文献[24]中比特币地址被认证，只要用户使用新的地址，就必须通过权威机构获取认证，大大降低了方案的执行效率。文献[25]设计了新的分布式匿名支付系统以解决监管问题，但系统只适用于 Zerocash[26]。本文提出了一种灵活的基于区块链的分布式 EHR 细粒度可追溯方案。

分布式密钥生成。DKG 协议是(t,n)阈值密码系统的组成部分之一[27]。它允许 n 方共同生成密钥

对（即公共密钥和私有密钥），而不需要让任何一方重建或存储密钥。如果不超过 $t+1$ 个参与方被破坏，则该协议是安全的。此外，文献[28]通过统一的随机性提高了 DKG 协议的安全性。通过运行 DKG 协议，每个诚实方将拥有密钥 a 的份额 a_i。对于 $\mathcal{N}$ 组中 $t+1$ 个正确份额，其中 $a=\sum_{i\in\mathcal{N}}\lambda_i a_i$，$\lambda_i$ 是集合 $\mathcal{N}$ 的拉格朗日内插系数，t−安全 DKG 协议将始终满足以下特性。

① 正确性。$t+1$ 份额的任何子集都定义相同的私钥 $a(a\in\mathbb{Z}_p)$，并且各方共享相同的公钥 $\mathrm{GP}=g^a$。

② 保密性。除了 $\mathrm{GP}=g^a$ 外，没有关于 a 的任何信息。

3　系统模型

系统模型如图 1 所示，涉及 7 个参与方：用户节点、医生节点、验证节点、仲裁节点、解密机构节点、区块链平台、分布式数据存储系统。表 1 列出了主要参数及含义。

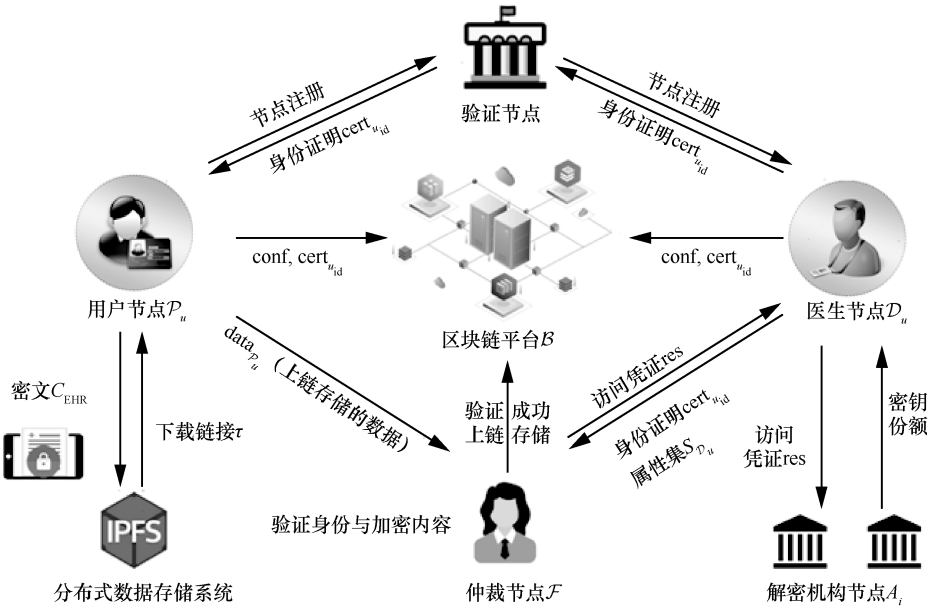

图 1　系统模型

表 1　　主要参数及含义

参数	含义
λ / MSK	安全参数/主密钥
pp/GP	公共参数/全局参数
A_i / PK_{A_i} / $\mathrm{PK}_{\mathrm{att}_j}$	解密机构/对应的公钥/属性公钥
u_{id} / regmess	用户节点的真实身份/节点的注册信息
(conf,pers)	节点的公开和保密信息对
$\mathrm{pk}_{\mathrm{chame}}$,$\mathrm{sk}_{\mathrm{chame}}$	变色龙哈希的公私钥对
$\mathrm{cert}_{u_{\mathrm{id}}}$	节点的身份证明
$(\boldsymbol{M},\rho)$	访问策略

验证节点主要用于生成链上节点的身份证明。用户生成注册信息，并将其发送给验证节点，完成注册信息的验证。成功注册的用户将得到由验证节点生成的身份证明，身份证明和私人信息之间存在一一对应的绑定关系。

用户节点由 $\mathcal{P}_u$ 标识，指可以提供 EHR 健康数据的患者。用户节点需要进行节点注册，从验证节点获得身份证明 $\mathrm{cert}_{u_{\mathrm{id}}}$。$\mathcal{P}_u$ 提交电子健康记录 EHR，用对称加密算法加密 EHR 得到 C_{EHR}，并发送到链外的 IPFS 中，IPFS 将返回一个下载链接 τ 用于检索数据，再用非对称加密算法将对称加密的密钥 ε 加密得到 C_ε。$\mathcal{P}_u$ 将 $\mathrm{data}_{\mathcal{P}_u}$ 传送给仲裁节点 $\mathcal{F}$。$\mathrm{data}_{\mathcal{P}_u}$ 经 $\mathcal{F}$ 验证后上传到区块链 $\mathcal{B}$ 存储，且 $\mathcal{P}_u$ 节点的公开信息和身份证明在链上公开显示。

医生节点由 $\mathcal{D}_u$ 标识，指申请访问 EHR 的实体。医生节点需要进行节点注册，从验证节点获得身份证明 $\mathrm{cert}_{u_{\mathrm{id}}}$，并将 $\mathrm{data}_{\mathcal{D}_u}$ 传送给仲裁节点 $\mathcal{F}$。$\mathrm{data}_{\mathcal{D}_u}$ 经 $\mathcal{F}$ 验证后上传到区块链 $\mathcal{B}$ 存储。当医生

节点 $\mathcal{D}_u$ 请求访问数据文件 $\text{data}_{\mathcal{D}_u}$ 时，$\mathcal{D}_u$ 将自己的属性集 $S_{\mathcal{D}_u}$ 以及身份证明 $\text{cert}_{u_{\text{id}}}$ 发送给链上仲裁节点 $\mathcal{F}$。验证通过后，$\mathcal{F}$ 返回访问凭证 res。$\mathcal{D}_u$ 获得访问凭证后，发送给解密机构节点 A_i，获得足够的密钥份额，获得 $\text{data}_{\mathcal{D}_u}$，完成解密操作。

仲裁节点由 $\mathcal{F}$ 标识，是指可以验证用户节点身份以及数据内容的节点，且可以为数据访问节点生成访问凭证。链上文件访问异常时，可通过查找用户身份实现追溯。

解密机构节点由 A_i 标识，是指使用 DKG 协议共同维护主密钥的实体。他们提供密钥份额以允许仲裁节点生成访问凭证 res。特别地，$\mathcal{D}_u$ 需要获得至少 $t+1$ 份额才能解密从区块链获得的数据。

区块链平台由 $\mathcal{B}$ 标识，是由多个区块链节点维护的许可区块链。上述几个角色都充当区块链节点，允许获得许可的参与者加入该区块链系统。追溯记录在 $\mathcal{B}$ 中，使公众可以审计调查的有效性和合法性。

分布式存储系统是指存储相关数据的数据存储系统，本文方案采用了分布式数据存储系统 IPFS。数据在加密后存入 IPFS 中，如果 $\mathcal{D}_u$ 想检索数据，则需要获得访问权限以获得解密密钥。

设定仲裁节点 $\mathcal{F}$ 和验证节点比其他节点具有更高的权重来维护 $\mathcal{B}$ 的安全性，区块链基于权益证明（PoS, proof of stake），如果存在可疑行为需要调查，$\mathcal{F}$ 将在区块链数据中追踪用户的真实身份，采用混合加密的方式加强数据保密性，提高加密效率。

4 算法设计

4.1 算法定义

1) 初始化算法

初始化算法输入安全参数 λ，得到公共参数 pp，主密钥 $\text{MSK}=a$，全局参数 $\text{GP}=g^a$。

2) 解密机构和用户节点初始化阶段

获得每个解密机构 A_i 的私钥为 SK_{A_i}，对应的公钥为 PK_{A_i} 以及属性公钥 PK_{att_j}。用户私钥 $\text{sk}_{G_{u_{\text{id}}}}$ 和公钥为 $\text{pk}_{G_{u_{\text{id}}}}$，且 $S_{A_{u_{\text{id}}}}$ 是解密机构对应用户 u_{id} 的属性集。用户 u_{id} 发送密钥份额申请 app 给解密机构 A_i，解密机构 A_i 将计算并返回属性私钥份额 SK_{att}。

3) 节点注册与节点验证阶段

节点注册阶段。输入公共参数 pp 和用户节点的真实身份 u_{id}，返回节点的注册信息 $\text{regmess}=(\text{CT}_{\text{mess}},\varsigma_{\text{mess}})$。

节点验证阶段。输入公共参数 pp、注册信息 regmess，以及注册信息的私钥 sk_{enrol}，验证节点身份 id 是否有效，验证有效，输出 1；否则输出 0。

4) 身份证明的生成和验证阶段

身份证明的生成阶段。输入公共参数 pp、节点的公开和保密信息对 (conf,pers)、用户身份的变色龙哈希值 $\text{CH}_{u_{\text{id}}}$、变色龙哈希的公私钥对 pk_{chame} 和 sk_{chame}、随机值 c、Merkle 树根节点 rt、树叶节点到树根节点的路径 $\text{path}_{u_{\text{id}}}$，输出节点的身份证明 $\text{cert}_{u_{\text{id}}}$。

身份证明的验证阶段。输入公共参数 pp、节点的公开信息 conf、节点的身份证明 $\text{cert}_{u_{\text{id}}}$，输出验证结果，若验证成功，则 $b=1$。

5) 数据生成阶段

在此阶段使用混合加密算法加密数据，且将数据密文存储于分布式存储系统 IPFS 中。对称加密算法，如 AES（advanced encryption standard）加密原数据，对称加密算法的密钥由 CP-ABE 实现加密。为了消除单点故障，本文方案采用 DKG 协议。假设有 $\mathcal{N}$ 个解密机构（也属于用户节点）分布于区块链上 $\{A_i\}_{i=1,2,\cdots,\mathcal{N}}$，每个解密机构 A_i 获得 a 的一个秘密份额 a_i，将数据加密的密文存储于 IPFS 中，IPFS 返回数据的下载链接。数据的密文 C_{EHR}、密文在 IPFS 中的下载链接 τ、数据加密的完成时间 timestamp，以及密钥加密结果都被上传到区块链中存储。

6) 数据访问权请求和授权阶段

当链上节点请求访问数据文件时，用户节点将自己的属性集合发送给链上仲裁节点 $\mathcal{F}$。如果用户的属性集合满足访问策略 $(\boldsymbol{M},\rho)$，则仲裁节点 $\mathcal{F}$ 将访问授权凭证发送给用户节点。

7) 解密并获取数据阶段

用户节点得到授权后，将获得凭证（包含足够的密钥份额），进而有权获得区块链中存储的 $\text{data}_{\mathcal{D}_u}$，可以完成解密操作。首先，用户节点利用访问授权凭证解密密钥密文，从而获得对称加密算法的密钥；然后，利用 IPFS 下载链接得到密文 C_{EHR}；最后，利用对称加密算法的密钥解密数据密文，从而得到数据明文，完成访问。

8) 追溯阶段

仲裁节点 $\mathcal{F}$ 获得链上数据 $\text{data}_{\mathcal{D}_u}$，解密后将得到其变色龙哈希公钥，通过检索验证节点上的记录信息，从而获得与用户节点变色龙哈希公钥对应的

用户真实 id，输出用户节点的身份序列 ID。

4.2　具体方案构造

1) 系统初始化阶段

$\text{Setup}(1^{\lambda})\to \text{pp}$。初始化算法输入安全参数 λ，得到公共参数 pp。公共参数组成如下：用于 zk-SNARK（zero knowledge succinct non-interactive argument of knowledge）中的 ς_{mess} 的证明和验证密钥对 $(\text{pk}_{\text{mess}},\text{vk}_{\text{mess}})$、用于 zk-SNARK 中 ς_{cert} 的证明和验证密钥对 $(\text{pk}_{\text{cert}},\text{vk}_{\text{cert}})$、注册机构的公钥 pk_{enrol} 以及用于变色龙哈希的公共参数 pp_{chame}。$\rho_{\text{nizk}}(\text{pk}_x,d_x,o_x)\to\varsigma_x$，其中，$x$ 为 mess 或 cert，表明给定 pk_x、陈述 d_x 和见证 o_x，ρ_{nizk} 将返回一个证明 ς_x。$o_{\text{nizk}}(\text{vk}_x,d_x,\varsigma_x)$ 用于验证，给定验证密钥 vk_x、陈述 d_x 和 ρ_{nizk} 所得证明 ξ_x，如果验证成功，则 o_{nizk} 返回 1，否则返回 0。

选择 $\mathbb{G}_1$ 和 $\mathbb{G}_2$ 是素数阶为 q 的循环群，且 $\mathbb{G}_1$ 的生成元为 g。设有双线性映射 $e:\mathbb{G}_1\times\mathbb{G}_1\to\mathbb{G}_2$、抗碰撞哈希函数 $H:\{0,1\}^*\to\mathbb{G}_1$。$\mathcal{N}$ 个解密机构（也属于用户节点）分布于区块链上 $\{A_i\}_{i=1,2,\cdots,\mathcal{N}}$，运行 DKG 协议，得到主密钥 $\text{MSK}=a$，其中 $a\in\mathbb{Z}_q$，每个解密机构 A_i 获得 a 的一个秘密份额 a_i。生成全局参数 $\text{GP}=g^a$。

2) 解密机构和用户节点初始化阶段

随机选取参数 $\eta_i,\theta_i,\kappa_i\in\mathbb{Z}_q$，每个解密机构 A_i 的私钥为 $\text{SK}_{A_i}=(\eta_i,\theta_i,\kappa_i)$，对应的公钥为 $\text{PK}_{A_i}=(e(g,g)^{\eta_i},g^{\frac{1}{\theta_i}},g^{\frac{\kappa_i}{\theta_i}})$。$A_i$ 随机选取 $x_j\in\mathbb{Z}_q$，生成属性公钥 $\text{PK}_{\text{att}_j}=(g^{x_j}H(j)^{\kappa_i})$。随机选取 $y,z\in\mathbb{Z}_q$，用户私钥 $\text{sk}_{G_{u_{\text{id}}}}=(y,z)$，公钥为 $\text{pk}_{G_{u_{\text{id}}}}=(g^y,g^z)$，且 $S_{A_{u_{\text{id}}}}$ 是解密机构对应用户 u_{id} 的属性集。u_{id} 发送密钥份额申请 app 给解密机构 A_i，如式(1)所示。

$$\text{app}=(g^{\frac{1}{z}},g^{ay},\text{PK}_{\text{att}_j}^{y}) \tag{1}$$

A_i 将计算并返回属性私钥份额，如式(2)所示。

$$\text{SK}_{\text{att}}=\left\{H=g^{\frac{\eta_i}{z}}g^{ay},I=g^{\frac{\theta_i}{z}},\ J=(g^{\frac{1}{z}})^{\theta_i\kappa_i}(g^{x_j}H(j)^{y\theta_j\kappa_j})\right\} \tag{2}$$

3) 节点注册与节点验证阶段

节点注册阶段可表示为

$$\text{Gen}_{\text{mess}}(\text{pp},u_{\text{id}})\to\text{regmess} \tag{3}$$

用户节点由 pp_{chame} 生成自己的变色龙方案的公钥和私钥对 $(\text{pk}_{\text{chame}},\text{sk}_{\text{chame}})$。由 u_{id} 计算其变色龙哈希值，随机选取 c，得到 $\text{CH}_{u_{\text{id}}}=H_{\text{chame}}\cdot(\text{pk}_{\text{chame}},u_{\text{id}},c)$。此时，用户可以为以下 NP 关系生成 zk-SNARK 证明 ς_{mess}。给定 $d_{\text{mess}}=(u_{\text{id}},\text{pk}_{\text{chame}},\text{CH}_{u_{\text{id}}})$，得到 $o_{\text{mess}}=(\text{sk}_{\text{chame}},c)$。其中，变色龙哈希的私钥与公钥的关系为 $\text{pk}_{\text{chame}}=\text{chame}_{\text{gen}}(\text{sk}_{\text{chame}})$。

仲裁节点 $\mathcal{F}$ 得到 $\varsigma_{\text{mess}}=\rho_{\text{nizk}}(\text{pk}_{\text{mess}},d_{\text{mess}},o_{\text{mess}})$。计算注册信息的加密密文 $\text{CT}_{\text{mess}}=\text{Enc}(\text{pk}_{\text{enrol}},d_{\text{mess}})$，验证节点存储 $(u_{\text{id}},\text{pk}_{\text{chame}},\text{sk}_{\text{chame}},c,\text{CH}_{u_{\text{id}}})$，同时返回

$$\text{regmess}=(\text{CT}_{\text{mess}},\varsigma_{\text{mess}}) \tag{4}$$

节点验证阶段可表示为

$$\text{Ver}_{\text{mess}}(\text{pp},\text{regmess},\text{sk}_{\text{enrol}})\to b \tag{5}$$

节点验证过程如图 2 所示，具体步骤如下。解析 regmess 为 $(\text{CT}_{\text{mess}},\varsigma_{\text{mess}})$，由解密算法得到 d_{mess}，$\text{Dec}(\text{sk}_{\text{enrol}},\text{CT}_{\text{mess}})=d_{\text{mess}}$，并解析 d_{mess} 为 $(u_{\text{id}},\text{pk}_{\text{chame}},\text{CH}_{u_{\text{id}}})$。如果所得 u_{id} 不是有效身份，则输出 b 值为 0；否则判断 $o_{\text{nizk}}(\text{vk}_{\text{mess}},d_{\text{mess}},\varsigma_{\text{mess}})=0$ 是否成立，若成立，则输出 $b=0$，否则 $b=1$。由仲裁节点存储 $(\text{pk}_{\text{chame}},u_{\text{id}},\text{CH}_{u_{\text{id}}})$，并通过 Merkle 树来发布 $(\text{pk}_{\text{chame}}\,\|\,\text{CH}_{u_{\text{id}}})$。

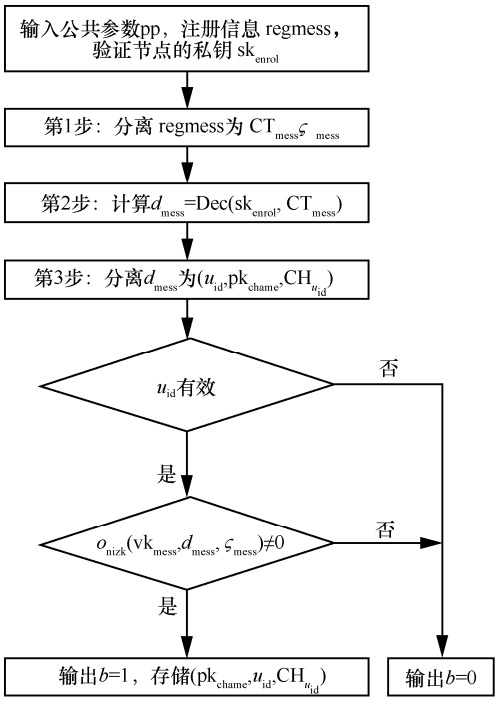

图 2　节点验证过程

4) 身份证明的生成和验证阶段

身份证明的生成阶段可表示为

$$\mathrm{Gen}_{\mathrm{cert}}(\mathrm{pp},(\mathrm{conf},\mathrm{pers}),\mathrm{CH}_{u_{\mathrm{id}}},(\mathrm{pk}_{\mathrm{chame}},\mathrm{sk}_{\mathrm{chame}}),c,\mathrm{rt},\mathrm{path}_{u_{\mathrm{id}}})\rightarrow \mathrm{cert}_{u_{\mathrm{id}}} \tag{6}$$

计算 $c'=\mathrm{CF}_{\mathrm{chame}}(\mathrm{sk}_{\mathrm{chame}},u_{\mathrm{id}},\mathrm{pers},c)$ 得到 c'，选取随机值 $\tilde{c}$ 用于加密。用户身份加密后为 $\mathrm{CT}_{u_{\mathrm{id}}}=\mathrm{Enc}(\mathrm{pk}_{\mathrm{enrol}},\mathrm{pk}_{\mathrm{chame}},\tilde{c})$。设置 2 个参数 w_{cert} 和 d_{cert}，$w_{\mathrm{cert}}=(\mathrm{rt},\mathrm{pk}_{\mathrm{enrol}},\mathrm{CT}_{u_{\mathrm{id}}})$，$d_{\mathrm{cert}}=(\mathrm{conf},w_{\mathrm{cert}})$。计算得到 o_{cert}，$o_{\mathrm{cert}}=(\mathrm{path}_{u_{\mathrm{id}}},\mathrm{CH}_{u_{\mathrm{id}}},\mathrm{sk}_{\mathrm{chame}},\mathrm{pk}_{\mathrm{chame}},\mathrm{pers},c',\tilde{c})$。计算 $\varsigma_{\mathrm{cert}}=\rho_{\mathrm{nizk}}(\mathrm{pk}_{\mathrm{cert}},d_{\mathrm{cert}},o_{\mathrm{cert}})$。最终输出用户节点的身份证明为

$$\mathrm{cert}_{u_{\mathrm{id}}}=(w_{\mathrm{cert}},\varsigma_{\mathrm{cert}}) \tag{7}$$

身份证明的验证阶段可表示为

$$\mathrm{Ver}_{\mathrm{cert}}=(\mathrm{pp},\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}}) \tag{8}$$

身份证明的验证过程如图 3 所示，主要用来验证用户节点的身份证明是否有效。分离 $\mathrm{cert}_{u_{\mathrm{id}}}$ 为 $(w_{\mathrm{cert}},\varsigma_{\mathrm{cert}})$，设置 $d_{\mathrm{cert}}=(\mathrm{conf},w_{\mathrm{cert}})$，若 $o_{\mathrm{nizk}}(\mathrm{vk}_{\mathrm{cert}},d_{\mathrm{cert}},\varsigma_{\mathrm{cert}})=0$，则返回 $b=0$，否则 $b=1$。

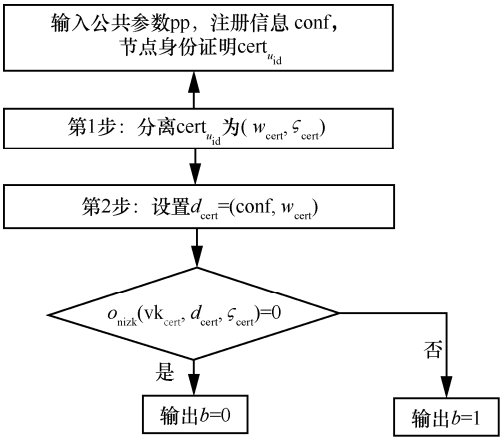

图 3 身份证明的验证过程

5) 数据生成阶段

假设患者 $\mathcal{P}_u$ 首先提交 EHR，用对称加密方法加密得到 C_{EHR} 并发送到链外的 IPFS 中，由 IPFS 返回下载地址 τ；然后用分布式 CP-ABE 来加密对称密钥 ε；最后，$\mathcal{P}_u$ 向区块链 $\mathcal{B}$ 提交一笔交易，将 $\mathrm{data}_{\mathcal{P}_u}=\{\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}},\mathrm{CT},\tau,C_{\varepsilon},\mathrm{timestamp}\}$ 存储到链上，可以使用区块链中的防篡改交易记录来验证数据的完整性。电子健康数据的生成存储过程具体如下。

① 数据加密

为了对 EHR 进行加密，随机生成 $\varepsilon\in\mathbb{G}_2$ 作为对称加密的密钥运行加密算法（例如 AES），对数据 EHR 进行加密，并计算 $\mathrm{Enc}_{\varepsilon}(\mathrm{EHR})=C_{\mathrm{EHR}}$。将 C_{EHR} 存储于 IPFS 中，IPFS 返回数据的下载链接 τ 给 $\mathcal{P}_u$。

② 对称加密算法中的密钥加密

患者 $\mathcal{P}_u$ 选择一个 $l\times n$ 的矩阵 $\boldsymbol{M}$，ρ 是 $\boldsymbol{M}$ 的行向量到用户属性的映射，$\boldsymbol{M}_{\bar{i}}$ 表示将 $\boldsymbol{M}$ 的第 $\bar{i}$ 行，$\bar{i}\in\{1,2,\cdots,l\}$。选取 $s\in\mathbb{Z}_q$ 为秘密值，在 $\mathbb{Z}_q$ 中随机选取 $\mu_2,\cdots,\mu_n$，设置向量 $\vec{v}=(s,\mu_2,\cdots,\mu_n)$。对于行向量 $\boldsymbol{M}_{\bar{i}}$，计算 $\xi_{\bar{i}}=\boldsymbol{M}_{\bar{i}}\vec{v}$。随机选取参数 $\varphi_1,\varphi_2,\cdots,\varphi_l$ 以及 $h_{\bar{i}}$，计算密钥 ε 的密文为

$$C_{\varepsilon}=(\tilde{C},\tilde{C}_{\bar{i},1},\tilde{C}_{\bar{i},2},\tilde{C}_{\bar{i},3}) \tag{9}$$

其中，

$$\tilde{C}=\varepsilon(\prod_{i=1}^{\mathcal{N}}e(g,g)^{\eta_i})^{s} \tag{10}$$

$$\tilde{C}_{\bar{i},1}=g^{a\xi_{\bar{i}}}(g^{v_{\rho(\bar{i})}}H(\rho(\bar{i})))^{h_{\bar{i}}\kappa_i} \tag{11}$$

$$\tilde{C}_{\bar{i},2}=g^{\frac{h_{\bar{i}}}{\theta_i}} \tag{12}$$

$$\tilde{C}_{\bar{i},3}=g^{\frac{h_{\bar{i}}\kappa_i}{\theta_i}} \tag{13}$$

$\mathcal{P}_u$ 公开信息 conf、身份证明 $\mathrm{cert}_{u_{\mathrm{id}}}$、身份 id 的加密结果 C、IPFS 中的密文下载地址 τ、对称密钥加密结果 C_{ε} 以及数据生成的时间戳 timestamp，并上传到区块链中存储，即

$$\mathrm{data}_{\mathcal{P}_u}=\{\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}},\mathrm{CT},\tau,C_{\varepsilon},\mathrm{timestamp}\} \tag{14}$$

6) 数据访问权请求和授权阶段

医生节点 $\mathcal{D}_u$ 将节点公开信息 conf、节点身份证明 $\mathrm{cert}_{u_{\mathrm{id}}}$、节点 id 加密结果 CT，以及数据生成时间戳 timestamp，并传送给仲裁节点 $\mathcal{F}$；$\mathrm{data}_{\mathcal{D}_u}=\{\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}},\mathrm{CT},\mathrm{timestamp}\}$ 经 $\mathcal{F}$ 验证后上传到区块链 $\mathcal{B}$ 存储。当链上医生节点 $\mathcal{D}_u$ 请求访问数据文件 $\mathrm{data}_{\mathcal{P}_u}$ 时，$\mathcal{D}_u$ 将自己的属性私钥发送给链上仲裁节点 $\mathcal{F}$。设置 $A_{\mathrm{set}_{u_{\mathrm{id}}}}$ 表示 $\mathcal{D}_u$ 的属性相关的所有解密机构，$A_{\mathrm{mun}_{u_{\mathrm{id}}}}$ 表示相关的解密机构的个数，$S_{\mathcal{D}_u}$ 表示医生用户 $\mathcal{D}_u$ 的所有属性集。$\mathcal{F}$ 选择集合 $\{k_j\in\mathbb{Z}_q\}_{j=1,2,\cdots,A_{\mathrm{num}_{u_{\mathrm{id}}}}}$，如果用户的属性集合 $S_{\mathcal{D}_u}$ 满足访问策略 $(\boldsymbol{M},\rho)$，则 $\sum_{i\in S_{\mathcal{D}_u}}k_j\lambda_j=s$，$\phi=g^{s}$。$\mathcal{F}$ 将访问授权凭证发送给 $\mathcal{D}_u$，凭证如下

$$\text{res}=\prod_{\tilde{i}\in A_{\text{set}_{u_{\text{id}}}}}(\theta_1\theta_2\theta_3) \tag{15}$$

其中，

$$\theta_1=\frac{e(\phi,s)}{\prod_{j\in S_{A_{u_{\text{id}}}}}(e(\tilde{C}_{\tilde{i},1},\text{pk}_{G_{u_{\text{id}}}}))^{k_j A_{\text{num}_{u_{\text{id}}}}}} \tag{16}$$

$$\theta_2=\frac{1}{\prod_{j\in S_{A_{u_{\text{id}}}}}(e(\tilde{C}_{\tilde{i},2},J_{u_{\text{id}}}))^{K_j A_{\text{num}_{u_{\text{id}}}}}} \tag{17}$$

$$\theta_3=\frac{1}{\prod_{j\in S_{A_{u_{\text{id}}}}}(e(\tilde{C}_{\tilde{i},3},I_{u_{\text{id}}}))^{k_j A_{\text{num}_{u_{\text{id}}}}}} \tag{18}$$

访问授权凭证为

$$\text{res}=\prod_{\tilde{i}\in A_{\text{set}_{u_{\text{id}}}}}(\theta_1\theta_2\theta_3)=\prod e(g,g)^{\frac{a_1 s}{z}} \tag{19}$$

7) 解密并获取数据阶段

$\mathcal{D}_u$ 得到授权后，将获得由 $\mathcal{F}$ 发送的访问权限凭证 res，从而获得区块链中存储的数据 $\text{data}_{\mathcal{F}_u}$，从足够数量的解密机构处获得的拥有密钥份额的解密凭证，从而可以完成解密操作。具体分为以下两步完成。

① $\mathcal{D}_u$ 利用访问授权凭证 res 和全局私钥 $\text{sk}_{G_{u_{\text{id}}}}$ 对 C_ε 解密，从而获得对称加密算法的密钥 ε。

②利用对称加密算法和解密密钥 ε 来解密 C_{EHR} 从而得到数据明文

$$\text{Dec}_\varepsilon(C_{\text{EHR}})=\text{EHR} \tag{20}$$

8) 追溯阶段

追溯过程如图 4 所示，可表示为

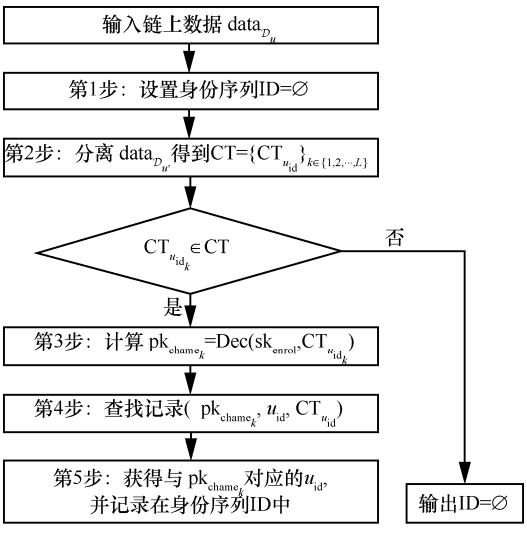

图 4　追溯过程

$$\text{Trace}(\text{data}_{\mathcal{F}_u})\rightarrow\text{ID} \tag{21}$$

初始化 ID =∅，获得链上数据 $\text{data}_{\mathcal{F}_u}$，解析得到用户节点身份加密结果 CT=$\{\text{CT}_{u_{\text{id}}}\}_{k\in\{1,2,\cdots,L\}}$，其中，$L$ 是区块链数据中医生节点的公开信息数量。每个 $\text{CT}_{u_{\text{id}_k}}\in\text{CT}$，解密得到 $\text{Dec}(\text{sk}_{\text{enrol}},\text{CT}_{u_{\text{id}_k}})=\text{pk}_{\text{chame}_k}$，检索链上记录 $(\text{pk}_{\text{chame}_k},u_{\text{id}},\text{CT}_{u_{\text{id}}})$，从而获得与 $\text{pk}_{\text{chame}_k}$ 对应的用户节点的身份 u_{id}。最终输出用户节点的身份序列 ID。

5　性能分析

本文方案与其他相关方案的功能对比如表 2 所示。分别对安全、隐私、用户认证、密钥管理以及可追溯性能进行比较。

表 2　本文方案与其他方案的功能对比

功能	安全	隐私	用户认证	密钥管理	可追溯
文献[15]方案	√	√	√	×	数据追溯
文献[29]方案	√	√	×	×	数据追溯
文献[30]方案	√	√	×	×	×
本文方案	√	√	√	DKG	ID 追溯

从表 2 可以看出，文献[15, 29-30]方案都满足安全与隐私性能，文献[15, 29]方案利用区块链的不变性追溯链上数据，但文献[15, 29-30]方案均缺乏密钥管理功能，存在安全威胁。比较结果表明，本文方案优于对比方案，可以为改善当前的分布式 EHR 应用提供解决方案。

5.1　设计实施

在区块链应用程序中通过私人信息生成公开信息的主要方式是使用椭圆曲线标量乘法，即 conf=pers$G$，其中，pers 是标量，G 是椭圆曲线的基点，而 conf 是椭圆曲线上的一个点。椭圆曲线标量乘法的实现基于 MNT6 椭圆曲线。使用 Java 编程语言基于 zk-SNARK 库 libsnark 实现本文方案的原型。在具体的实现中，使用提供 128 bit 安全性的 Barreto-Naehrig 椭圆曲线作为 zk-SNARK 方案的基础曲线。变色龙哈希和公钥加密方案的实现基于 254 bit 的质数字段。基于 JPBC（基于 PBC 配对的密码库）在 DKG 协议和 CP-ABE 上实现密码算法，采用具有 160 bit $\mathbb{Z}_q$ 和 512 bit $\mathbb{G}_1$ 的素数双线性群设置系统参数。

本节实验在个人计算机上运行，主要参数是 IntelCore™ i5-5200U、2.20 GHz CPU，以及 4 GB

RAM。基于 5 个授权机构（5 个中的 3 个）设计 DKG 协议，以评估通信性能和时间开销。设置阶段包括 DKG 协议和 CP-ABE 的初始化，使用混合加密方法计算加密和解密的时间性能。

5.2 时间开销与通信开销

设置 Merkle 树支持的最大用户数分别为 2^{10}、2^{20}、2^{30}，评估算法性能。图 5 和图 6 展示了 Setup、Gen_{mess}、Ver_{mess}、Gen_{cert}、Ver_{cert} 和 Trace 算法的时间开销。其中，time 代表算法的运行时间。|·|代表数据长度。例如 $|\text{vk}_{\text{mess}}|$ 表示注册中用于 zk-SNARK 中验证密钥的长度。

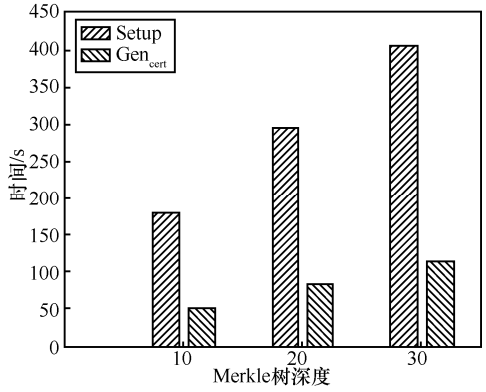

图 5 Setup 和 Gen_{cert} 算法的时间开销

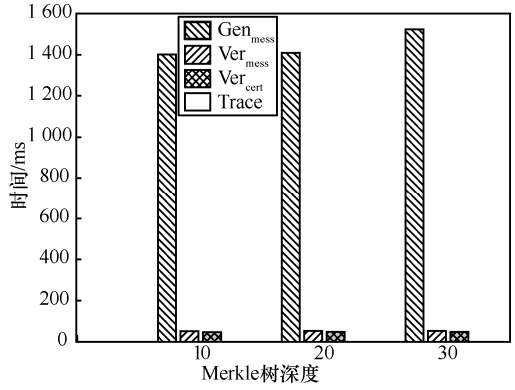

图 6 Gen_{mess}、Ver_{mess}、Ver_{cert} 和 Trace 算法的时间开销

1) Setup 算法在 Merkle 树深度为 10 时，运行时间为 180 s；深度为 20 时，运行时间为 295 s；深度为 30 时，运行时间为 401 s。

2) Gen_{mess} 算法在 Merkle 树深度为 10 时，运行时间为 1 401 ms；深度为 20 时，运行时间为 1 409 ms；深度为 30 时，运行时间为 1525 ms。

3) Ver_{mess} 算法在 Merkle 树深度为 10 时，运行时间为 50.7 ms；深度为 20 时，运行时间为 52.6 ms；深度为 30 时，运行时间为 53.3 ms。

4) Gen_{cert} 算法在 Merkle 树深度为 10 时，运行时间为 53 s；深度为 20 时，运行时间为 81 s；深度为 30 时，运行时间为 113 s。

5) Ver_{cert} 算法在 Merkle 树深度为 10 时，运行时间为 46.9 ms；深度为 20 时，运行时间为 48.1 ms；深度为 30 时，运行时间为 48.3 ms。

6) Trace 算法运行时间与 Merkle 树深度无关，均为 0.17 ms。

生成证明和验证密钥对的通信开销如表 3 所示。由表 3 知，用于 zk-SNARK 中的 ς_{mess} 的证明和验证密钥对 $(\text{pk}_{\text{mess}},\text{vk}_{\text{mess}})$ 的大小分别为 663 KB 和 665 B。在树深度为 30 时，用于 zk-SNARK 中的 ς_{cert} 的证明和验证密钥对 $(\text{pk}_{\text{cert}},\text{vk}_{\text{cert}})$ 的大小分别为 237 MB 和 11 KB。zk-SNARK 中 ς_{mess} 和 ς_{cert} 的大小为 341 B。

设 S 为方案加密健康数据文件时选取的属性集，S_{sk_u} 为与用户属性私钥相关的属性集，t_{e_1}、t_{e_2} 分别为在群 $\mathbb{G}_1$ 和 $\mathbb{G}_2$ 上进行模幂运算消耗的时间，t_p 是双线性配对操作消耗的时间。数据文件的加密解密与属性私钥和全局私钥相关，能够有效避免解密机构节点 A_i 之间的联合攻击。EHR 加密与解密时间开销如图 7 所示。由图 7 可知，方案加密数据的时间开销与加密健康数据文件时选取的属性集空间规模 $|S|$ 成正相关，表达式为 $4|S|t_{e_1}+t_{e_2}$。而解密数据的时间开销与属性无关，表达式为 t_{e_2}。

设 $|\mathbb{G}_1|$ 和 $|\mathbb{G}_2|$ 分别为群 $\mathbb{G}_1$ 和 $\mathbb{G}_2$ 的空间规模，$|S_{\text{sk}_u}|$ 为与用户属性私钥相关的属性集的空间规模。EHR 数据加密与解密通信开销如表 4 所示。

表 3 生成证明和验证密钥对的通信开销

Merkle 树深度	pk_{mess} /KB	vk_{mess} /B	pk_{cert} /MB	vk_{cert} /KB	ς_{cert} /B	ς_{mess} /B
10	663	665	103	11	341	341
20	663	665	171	11	341	341
30	663	665	237	11	341	341

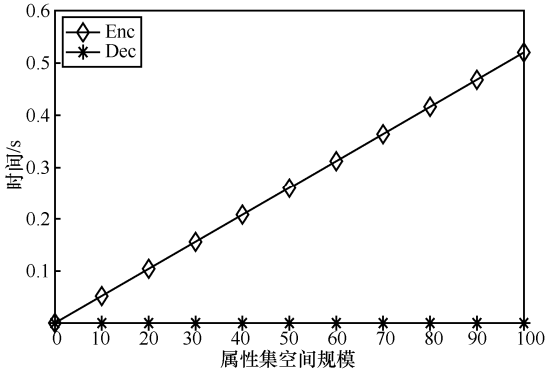

图 7　EHR 加密数据与解密数据时间开销

表 4　EHR 加密数据与解密数据通信开销

阶段	通信开销
用户节点数据加密	$3\|S\|\times\|\mathbb{G}_1\|+\|\mathbb{G}_2\|$
访问授权与解密	$3\|S_{\mathrm{sk}_u}\|\times\|\mathbb{G}_1\|$

由表 4 可知，用户节点数据加密的通信开销与加密健康数据文件时选取的属性集的空间规模$|S|$成正相关，与群$|\mathbb{G}_1|$和$|\mathbb{G}_2|$均相关，相应表达式为$3|S|\times|\mathbb{G}_1|+|\mathbb{G}_2|$。而访问授权与解密数据的通信开销与用户属性私钥相关的属性集$|S_{\mathrm{sk}_u}|$成正相关，仅与群$|\mathbb{G}_1|$有关，表达式为$3|S_{\mathrm{sk}_u}|\times|\mathbb{G}_1|$。

6　安全性证明

追溯算法的安全性证明主要通过节点身份证明的不可区分性来实现。攻击游戏双方分别是敌手$\mathcal{A}$和挑战者$\mathcal{C}$，具有身份 ID 的诚实用户的行为由预言机$\mathcal{O}_{\mathrm{id}}$实现。假设实验中的诚实用户和敌手已经在注册认证机构中成功注册，可以生成任何身份证明。挑战者$\mathcal{C}$使用公共参数 pp 初始化$\mathcal{O}_{\mathrm{id}}$。

$\mathcal{O}_{\mathrm{id}}$存储用于生成注册信息的秘密信息$\mathrm{pers}_{\mathrm{enrol}}$、用户生成的一组身份证明$\mathrm{cert}_{u_{\mathrm{id}}}$以及用户用来生成身份证明的一组证据$\mathrm{cert}_{\mathrm{pers}}$。$\mathcal{O}_{\mathrm{id}}$接受 2 种不同的查询，具体如下。

1) 查询 1。敌手不知道私人信息 pers。$\mathcal{O}_{\mathrm{id}}$首先随机选择 pers，生成公共信息 conf 发布；然后，$\mathcal{O}_{\mathrm{id}}$调用$\mathrm{Gen}_{\mathrm{cert}}$算法以生成身份证明$\mathrm{cert}_{u_{\mathrm{id}}}$，并将$(\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}})$发送给查询者。

2) 查询 2。敌手知道私人信息 pers，$\mathcal{O}_{\mathrm{id}}$首先使用 pers 生成公共信息 conf 发布，然后调用$\mathrm{Gen}_{\mathrm{cert}}$算法来生成身份证明$\mathrm{cert}_{u_{\mathrm{id}}}$，$\mathcal{O}_{\mathrm{id}}$将$(\mathrm{conf},\mathrm{cert}_{u_{\mathrm{id}}})$发送给查询者。

攻击游戏 1　$\mathrm{Game}_{\mathrm{inden}}$：身份证明不可区分

步骤 1　挑战者$\mathcal{C}$选取随机值$b\in\{0,1\}$，通过运行$\mathrm{Setup}(1^{\lambda})$获得公共参数 pp，并发送给敌手$\mathcal{A}$。$\mathcal{C}$初始化 2 个独立的预言机$\mathcal{O}_{\mathrm{id}_0}$和$\mathcal{O}_{\mathrm{id}_1}$。

步骤 2　敌手$\mathcal{A}$发送一系列查询。

步骤 3　挑战者$\mathcal{C}$对于$\mathcal{A}$的查询回复为$(\mathrm{conf}_b,\mathrm{cert}_{u_{\mathrm{id}_b}})$。

步骤 4　查询结束，$\mathcal{A}$向$\mathcal{C}$发送一个猜测$b'\in\{0,1\}$。如果$b=b'$，$\mathcal{C}$输出 1，挑战成功；否则，$\mathcal{C}$输出 0。

身份证明的不可区分性要求敌手$\mathcal{A}$以极低的概率赢得上述游戏。正式定义如定义 1 所示。

定义 1　如果对于任意多项式时间敌手$\mathcal{A}$，方案都满足身份证明的不可区分性，则敌手$\mathcal{A}$存在可忽略优势$\varepsilon_{\mathrm{inden}}$，使

$$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{inden}}=\Pr[b'=b]-\frac{1}{2}\leqslant\varepsilon_{\mathrm{inden}}$$

定理 1　假设完全为零知识，加密方案 Enc 满足 IND-CCA2 安全性，则本方案满足身份证明不可区分性。

证明　下面通过一系列混合游戏证明定理 1。令Q_m为敌手$\mathcal{A}$发送的查询数。定义一组游戏如下。

$\mathrm{Game}_{\mathrm{real}}$表示真实的 IND-CCA2 游戏设置。

Game_0与$\mathrm{Game}_{\mathrm{real}}$除了挑战者$\mathcal{C}$将模拟方案之外，其他设置相同。挑战者$\mathcal{C}$调用多项式时间模拟器$S_{\mathrm{nizk}}(\lambda,\mathrm{AC}_{\mathrm{cert}})$获得证明和验证密钥对$(\mathrm{pk}_{\mathrm{cert}},\mathrm{vk}_{\mathrm{cert}})$以及门限 trap。由于本文方案完全是零知识，因此模拟$\varsigma_{\mathrm{cert}}$的分布与证明在$\mathrm{Game}_{\mathrm{real}}$中计算。因此，$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_{\mathrm{real}}}=\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_0}$。

Game_1与Game_0设置相同，除了挑战者$\mathcal{C}$替换了$C_{u_{\mathrm{id}}}$，Game_1中通过加密一个随机字符串来生成$\mathrm{cert}_{u_{\mathrm{id}}}$。当预言机$\mathcal{O}_{\mathrm{id}}$发送身份证$\mathrm{cert}_{u_{\mathrm{id}}}$给挑战者$\mathcal{C}$时，$\mathcal{C}$替换$C_{u_{\mathrm{id}}}$为$C'_{u_{\mathrm{id}}}$，其中，$C'_{u_{\mathrm{id}}}$由算法$\mathrm{Enc}(\mathrm{pk}_{\mathrm{enrol}},\mathrm{pk}_{\mathrm{chame}},\tilde{r}')$生成，$\tilde{r}'$是在身份加密方案中的明文空间均匀采样的随机字符串。在Game_1中，敌手$\mathcal{A}$与b无关，所以$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_1}=0$。

只需要证明不存在多项式时间的敌手$\mathcal{A}$能够以可忽略的概率区分Game_0和Game_1。

构造算法$\mathcal{C}'$，且$\mathcal{A}$作为一个子程序赢得 IND-CCA2 游戏。定义$\varepsilon_{\mathrm{IND}}=\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_1}-\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_0}$。

当 $\mathcal{A}$ 发出第 i 个查询时，$i\in\{1,2,\cdots,Q_{u_m}\}$，$\mathcal{C}'$ 提取 $\varepsilon_{\mathrm{IND}}$ 获得与 $C_{u_{\mathrm{id}}}$ 对应的明文 m'。然后，$\mathcal{C}'$ 随机选取字符串 $\tilde{r}$，$\tilde{r}$ 与明文 m' 等长。$\mathcal{C}'$ 发送 $(m_0,m_1)=(m',\tilde{r})$ 给挑战者 $\mathcal{C}$，$\mathcal{C}$ 生成 $\mathrm{Enc}(\mathrm{pk}_{\mathrm{enrol}},m_b,\tilde{r})$ 并返回给 $\mathcal{C}'$。$\mathcal{C}'$ 替换 $C_{u_{\mathrm{id}}}$ 为 $C'_{u_{\mathrm{id}}}$。查询结束，$\mathcal{A}$ 发送猜测 $b'\in\{0,1\}$。如果 $b=b'$，$\mathcal{C}$ 输出 1，挑战成功；否则，$\mathcal{C}$ 输出 0。

基于对每个查询密文的标准混合参数，可以得出结论，在实验的随机性上，$\mathcal{C}'$ 必须以至少 $\varepsilon_{\mathrm{IND}}/Q_{u_m}$ 的优势在 IND-CCA2 实验中成功。因此，

$$\left|\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_1}-\mathrm{Adv}_{\mathcal{A}}^{\mathrm{Game}_0}\right|\leqslant\varepsilon_{\mathrm{IND}}。$$

证毕。

定理 2 如果 q-BDHE 假设成立，则方案在随机预言机模型下是自适应安全的。

证明 假设敌手任意选择访问结构 $(\boldsymbol{M}',\rho')$，其中 $\boldsymbol{M}'$ 的列数为 n'，且 $n'<q$。

询问预言机 $\mathcal{O}_{\mathrm{dkg}}$，挑战者 $\mathcal{C}$ 响应敌手的询问。

步骤 1 敌手 $\mathcal{A}$ 发起对解密机构 A'_i 的询问，挑战者 $\mathcal{C}$ 选取随机参数 $\eta'_i,\theta'_i,\kappa'_i\in\mathbb{Z}_q$，并将对应公钥 $\mathrm{PK}_{A'_i}=(e(g,g)^{\eta'_i},g^{1/\theta'},g^{\kappa'_i/\theta'_i})$ 发送给 $\mathcal{A}$。

步骤 2 敌手 $\mathcal{A}$ 询问属性 att 公钥，挑战者 $\mathcal{C}$ 随机选取 $x'_i\in\mathbb{Z}_q$，生成属性公钥 $\mathrm{PK}_{\mathrm{att}'_j}=(g^{x'_j}H(att'_j))^{\kappa'_i}$ 并发送给 $\mathcal{A}$。

步骤 3 挑战 $\mathcal{C}$ 随机选取参数 $r'\in\mathbb{Z}_q$，向量 $\vec{w}=(w_1,w_2,\cdots,w_{n'})\in\mathbb{Z}_q$，且 $\mathrm{pk}_{G_{u_{\mathrm{id}}}}=g^z$。

步骤 4 敌手 $\mathcal{A}$ 将 2 个等长的消息 m_0 和 m_1 发送给挑战者 $\mathcal{C}$，$\mathcal{C}$ 选取随机值 $b\in\{0,1\}$，并且生成 $\tilde{C}=m_bT(\prod_{i=1}^{\mathcal{N}}e(g,g)^{\eta_i})^s$ 和 $\phi'=g^s$。$\mathcal{C}$ 在 $\mathbb{Z}_q$ 中随机选取 $\mu'_2,\cdots,\mu'_n$ 和 $\delta'_1,\delta'_2,\cdots,\delta'_l$。秘密值 s 通过向量 $\vec{v}$ 分享，其中 $\vec{v}=(s,sa+\mu'_2,sa^2+\mu'_3,\cdots,sa^{n'-1}+\mu'_{n'})$。计算 $\tilde{C}_{\bar{i},1}=g^{a\xi_{\bar{i}}}(g^{v_{\rho(\bar{i})}}H(\rho(\bar{i})))^{h_{\bar{i}}\kappa_i}$，$\tilde{C}_{\bar{i},2}=g^{h_{\bar{i}}/\theta_i}$，$\tilde{C}_{\bar{i},3}=g^{h_{\bar{i}}\kappa_i/\theta_i}$，将 $(\tilde{C},\tilde{C}_{\bar{i},1},\tilde{C}_{\bar{i},2},\tilde{C}_{\bar{i},3})$ 发送给 $\mathcal{A}$。

步骤 5 查询结束，$\mathcal{A}$ 向 $\mathcal{C}$ 发送一个猜测 $b'\in\{0,1\}$。如果 $b=b'$，$\mathcal{C}$ 输出 1，$\mathcal{A}$ 破解 q-BDHE 困难问题的优势为 $\mathrm{Adv}_{\mathcal{A}}^{q-\mathrm{BDHE}}=\Pr[b'=b]-1/2\leqslant\varepsilon_{q-\mathrm{BDHE}}$，挑战成功，$T=e(g,g)^{sa^{q+1}}$；否则，$\mathcal{C}$ 输出 0，T 是群 $\mathbb{G}_2$ 中的随机参数，敌手 $\mathcal{A}$ 没有获得有价值的信息。以上证明可知，方案在随机预言机模型下是自适应安全的。证毕。

7 结束语

本文针对 EHR 在分布式存储中的密钥管理及用户身份追溯问题，提出了一种基于区块链的分布式 EHR 细粒度可追溯方案；结合变色龙哈希和 zk-SNARK 证明技术实现了链上节点的注册与身份证明的生成，从而能够对链上用户的追溯。此外，本文通过设置多个解密机构节点联合分发用户节点的属性私钥，消除了分布式系统中密钥管理的单点故障问题，设计的分布式密文策略的属性基加密方案实现了安全细粒度的访问控制。安全性证明表明，本文方案在随机预言机模型下是自适应安全的。理论分析和实验结果表明了本文方案在电子健康数据传输和共享方面是可行且实用的。

参考文献：

[1] OVERGAARD S M. Harnessing the power of data in health[R]. Palo Alto: Stanford University, 2017.

[2] ZYSKIND G, NATHAN O, PENTLAND A. Decentralizing privacy: using blockchain to protect personal data[C]//2015 IEEE Security and Privacy Workshops. Piscataway: IEEE Press, 2015: 180-184.

[3] ZAGHLOUL E, LI T T, MUTKA M W, et al. Bitcoin and blockchain: security and privacy[J]. IEEE Internet of Things Journal, 2020, 7(10): 10288-10313.

[4] PANDEY P, LITORIYA R. Implementing healthcare services on a large scale: challenges and remedies based on blockchain technology[J]. Health Policy and Technology, 2020, 9(1): 69-78.

[5] JIANG J X, BAI G. Evaluation of causes of protected health information breaches[J]. JAMA Internal Medicine, 2019, 179(2): 265-267.

[6] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system[R]. 2008.

[7] FENG Q, HE D B, ZEADALLY S, et al. A survey on privacy protection in blockchain system[J]. Journal of Network and Computer Applications, 2019, 126: 45-58.

[8] 王明生, 曹鹤阳, 李佩瑶. 基于区块链的去中心化信贷系统及应用[J]. 通信学报, 2019, 40(8): 169-177.
WANG M S, CAO H Y, LI P Y. Decentralized credit system based on blockchain and its application[J]. Journal on Communications, 2019, 40(8): 169-177.

[9] XU J, XUE K P, LI S H, et al. Healthchain: a blockchain-based privacy preserving scheme for large-scale health data[J]. IEEE Internet of Things Journal, 2019, 6(5): 8770-8781.

[10] 熊金波, 毕仁万, 陈前昕, 等. 边缘协作的轻量级安全区域建议网络[J]. 通信学报, 2020, 41(10): 188-201.
XIONG J B, BI R W, CHEN Q X, et al. Towards edge-collaborative, lightweight and secure region proposal network[J]. Journal on Communications, 2020, 41(10): 188-201.

[11] 史锦山, 李茹. 物联网下的区块链访问控制综述[J]. 软件学报,

2019, 30(6): 1632-1648.
SHI J S, LI R. Survey of blockchain access control in Internet of Things[J]. Journal of Software, 2019, 30(6): 1632-1648.

[12] LIANG W, LONG J, WENG T H, et al. TBRS: a trust based recommendation scheme for vehicular CPS network[J]. Future Generation Computer Systems, 2019, 92: 383-398.

[13] AZARIA A, EKBLAW A, VIEIRA T, et al. MedRec: using blockchain for medical data access and permission management[C]//2016 2nd International Conference on Open and Big Data. Piscataway: IEEE Press, 2016: 25-30.

[14] DAGHER G G, MOHLER J, MILOJKOVIC M, et al. Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology[J]. Sustainable Cities and Society, 2018, 39: 283-297.

[15] XIA Q, SIFAH E, SMAHI A, et al. BBDS: blockchain-based data sharing for electronic medical records in cloud environments[J]. Information, 2017, 8(2): 44.

[16] FAN K, WANG S Y, REN Y H, et al. MedBlock: efficient and secure medical data sharing via blockchain[J]. Journal of Medical Systems, 2018, 42(8): 1-11.

[17] HUSSEIN A F, ARUNKUMAR N, RAMIREZ-GONZALEZ G, et al. A medical records managing and securing blockchain based system supported by a Genetic Algorithm and Discrete Wavelet Transform[J]. Cognitive Systems Research, 2018, 52: 1-11.

[18] 闫玺玺, 原笑含, 汤永利, 等. 基于区块链且支持验证的属性基搜索加密方案[J]. 通信学报, 2020, 41(2): 187-198.
YAN X X, YUAN X H, TANG Y L, et al. Verifiable attribute-based searchable encryption scheme based on blockchain[J]. Journal on Communications, 2020, 41(2): 187-198.

[19] WANG G J, LIU Q, WU J. Hierarchical attribute-based encryption for fine-grained access control in cloud storage services[C]//The 17th ACM Conference on Computer and Communications Security. New York: ACM Press, 2010: 735-737.

[20] PASUPULETI S K, ALPHONSE P J A, PREMKAMAL P K. Efficient revocable CP-ABE for big data access control in cloud computing[J]. International Journal of Security and Networks, 2019, 14(3): 119.

[21] LEWKO A, WATERS B. Decentralizing attribute-based encryption[M]. Berlin: Springer, 2011.

[22] HU S, LI J, ZHANG Y. Improving security and privacy-preserving in multi-authorities ciphertext-policy attribute-based encryption[J]. KSII Transactions on Internet and Information Systems, 2018, 12(10): 5100-5119.

[23] LI J G, HU S Z, ZHANG Y C. Two-party attribute-based key agreement protocol with constant-size ciphertext and key[J]. Security and Communication Networks, 2018, 2018: 1-10.

[24] ATENIESE G, FAONIO A, MAGRI B, et al. Certified bitcoins[M]. Cham: Springer International Publishing, 2014.

[25] GARMAN C, GREEN M, MIERS I. Accountable privacy for decentralized anonymous payments[C]//International Conference on Financial Cryptography and Data Security. Berlin: Springer, 2016: 81-98.

[26] BEN S E, CHIESA A, GARMAN C, et al. Zerocash: decentralized anonymous payments from bitcoin[C]//2014 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2014: 459-474.

[27] PEDERSEN T P. Non-interactive and information-theoretic secure verifiable secret sharing[C]//Annual International Cryptology Conference. Berlin: Springer, 1991: 129-140.

[28] GENNARO R, JARECKI S, KRAWCZYK H, et al. Secure distributed key generation for discrete-log based cryptosystems[C]//International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 1999: 295-310.

[29] RAMANI V, KUMAR T, BRACKEN A, et al. Secure and efficient data accessibility in blockchain based healthcare systems[C]//2018 IEEE Global Communications Conference. Piscataway: IEEE Press, 2018: 206-212.

[30] NGUYEN D C, PATHIRANA P N, DING M, et al. Blockchain for secure EHRs sharing of mobile cloud based E-health systems[J]. IEEE Access, 2019, 7: 66792-66806.

[作者简介]

应作斌（1982− ），男，安徽芜湖人，博士，安徽大学讲师，主要研究方向为云安全、应用密码学等。

斯元平（1994− ），女，安徽安庆人，安徽大学硕士生，主要研究方向为应用密码学、区块链技术等。

马建峰（1963− ），男，陕西西安人，博士，西安电子科技大学教授、博士生导师，主要研究方向为计算机系统安全、移动与无线安全、系统可生存性和可信计算。

刘西蒙（1988− ），男，陕西西安人，博士，福州大学教授，主要研究方向为云安全、应用密码学、大数据安全等。