第46卷 第4期 计 算 机 工 程 2020年4月
Vol. 46 No. 4 Computer Engineering April 2020

·网络空间安全· 文章编号: 1000-3428(2020)04-0143-08 文献标志码: A 中图分类号: TP309

无证书的共享数据公开审计方案

刘雪艳,贺啸梅,芦婷婷,罗玉坤

(西北师范大学 数学与统计学院,兰州 730070)

摘 要: 公钥密码体制中多数公开审计方案存在证书管理问题,会增加存储负荷和通信成本。为有效验证半可信云中数据的完整性,减少证书管理的额外开销,提出一种无证书的公开审计方案。采用同态技术实现批审计,高效完成多个用户的审计需求,通过ELGamal加密体制对用户身份进行追踪,防止用户的恶意行为。安全性和性能分析结果表明,该方案安全高效,能够抵抗类型Ⅰ和类型Ⅱ敌手攻击,并满足签名不可伪造性和签名用户身份隐私性。

关键词: 无证书方案; 公开审计; 批审计; 隐私保护; 不可伪造性

开放科学(资源服务)标志码(OSID):

中文引用格式: 刘雪艳,贺啸梅,芦婷婷,等. 无证书的共享数据公开审计方案[J]. 计算机工程,2020,46(4):143-150.

英文引用格式: LIU Xueyan,HE Xiaomei,LU Tingting,et al. Certificateless public audit scheme for shared data[J]. Computer Engineering,2020,46(4):143-150.

Certificateless Public Audit Scheme for Shared Data

LIU Xueyan,HE Xiaomei,LU Tingting,LUO Yukun

(College of Mathematics and Statistics,Northwest Normal University,Lanzhou 730070,China)

【Abstract】 Many public audit schemes in public key cryptosystem have certificate management problem,which will increase storage load and communication cost. In order to effectively verify the integrity of the data in the semi-trusted cloud and reduce the extra cost of certificate management,this paper proposes a certificateless public audit scheme. The batch audit is realized by using homomorphic technology,so as to efficiently complete the audit needs of multiple users. The ELGamal encryption system is adopted to track user identity,thus preventing the malicious behavior of users. The results of security and performance analysis show that the proposed scheme is safe and efficient. It can resist type Ⅰ and type Ⅱ adversary attacks and satisfy the unforgeability of signature and the privacy of user identity.

【Key words】 certificateless scheme; public audit; batch audit; privacy protection; unforgeability

DOI: 10.19678/j.issn.1000-3428.0054698

0 概述

随着计算机技术的快速发展和网络数据的海量增加,云存储技术成为云计算不可或缺的一部分,它允许用户把大量的数据外包存储在云中,从而减少用户存储空间管理和计算的成本,但此时外包数据的安全也受到极大威胁,不可信的云存储提供者可能为了利益而篡改或删除数据,也有可能因为硬件/软件故障而造成数据丢失。因此,用户需要确认存储在云中的数据是未被篡改且被云存储器完整存储的[1]。

为有效验证外包数据的完整性,研究者相继提出很多审计方案。当用户想要检验数据的完整性时,第三方审计者(Third Party Auditor,TPA)能够代替用户在不下载所有数据的前提下对数据进行完整性的检验。如果TPA在执行审计过程中没有删除用户的任何秘密值,则称为公开审计。文献[2]提出公开审计的方案,而为提高性能和安全性,文献[3-5]提出了新的审计方案,文献[3]方案支持无需第三方审计者帮助的公开审计,且不会向第三方审计者泄露用户的隐私信息,文献[4]提出一种灵活的分布式存储完整性审计机制,文献[5]提出一种安全的云存储系统支持保护隐私的公开审计方案。此外,文献[6-8]方案则是在前人所做工作基础上的改进。文献[6]针对半可信的第三方审计者,提出一种用户可以与云服

基金项目: 国家自然科学基金(61662071,61562077)。

作者简介: 刘雪艳(1978—),女,副教授,主研方向为属性密码学、信息安全; 贺啸梅、芦婷婷、罗玉坤,硕士研究生。

收稿日期: 2019-04-23 修回日期: 2019-06-16 E-mail: liuxy@nwnu.edu.cn

务提供者进行交互，由用户自己完成数据完整性验证的方案，文献[7]提出一种改进的基于 LBT 树形认证结构的数据完整性公开审计方案，文献[8]提出已知数据伪造改造的概念，利用基于等级的认证跳表设计了相应的改进方案。之后，针对动态群组隐私保护的审计方案[9-10]和面向公有云及多管理者群组的公开审计方案[11]相继被提出。上述审计方案都是基于传统的公钥密码体制，由证书授权方产生用户的公私钥对，通过绑定用户的公钥来产生对应的私钥，而产生的证书都是由证书授权方管理，会出现证书管理问题，包括密钥的分发、存储、撤销和认证，并且密钥生成中心(Key Generation Center，KGC)有能力产生任何实体的签名，有可能会泄露用户的身份信息。针对证书管理问题，一系列基于身份的公开审计方案[12-14]被提出。文献[12]提出一种新的基于身份的聚合签名公开审计方案，解决了存储用户在上传数据前需要先颁发证书从而产生巨大成本的问题，并在其安全模型下是可证明安全的，文献[13]采用一种有效的基于证书的公钥设置密钥管理方案，将基于身份的聚合签名与公开验证相结合，构造了可证明数据完整性协议，减少了 TPA 单个任务的审计时间，文献[14]则提出针对多个云环境的基于身份的审计方案。

然而，上述方案都存在密钥托管问题，用户的私钥完全是由 KGC 产生，而 KGC 有可能不是完全可信的，可能会泄露用户的隐私信息。在这种情况下，该问题可以通过设定 KGC 是完全可信的实体而得以解决，这在文献[15-16]的方案中有所体现。文献[17]提出了无证书的公开审计(Certificateless Public Key Cryptography，CLPKC)的概念。在 CLPKC 中，用户的私钥包含两部分，一部分是由用户自己生成的，另一部分是由 KGC 生成的。因此，CLPKC 解决了传统公钥密码体制中的证书管理和密钥托管问题。基于先前的工作，文献[18]提出了新的签名机制——无证书签名机制，该机制由无块验证和同态认证来实现，是首个通过无证书签名产生的无证书公开审计方案。该方案的提出解决了传统密码体制中的证书管理问题和基于身份公开审计中的密钥托管问题，但其签名方案不能抵抗第一类型攻击，即攻击者可以在他的意愿下替换用户的公钥，并且当用户出现恶意行为时，不支持用户追踪的功能。

本文提出一种无证书的公开审计方案，采用同态技术完成多用户的审计请求，从而减少计算量，提高审计效率。同时利用 ELGamal 体制对恶意用户身份进行追踪，以保证数据和签名用户身份的隐私性。

1 系统模型和安全模型

1.1 系统模型

本文方案的系统模型如图 1 所示，其中包含4 个实体：KGC，TPA，云存储器提供者(Cloud Server Provider，CSP)，群组用户(Users)。每类实体的主要功能描述如下：

1) KGC：KGC 是完全可信的，它生成系统主密钥、公共参数以及产生用户和群管理者的部分公私钥对。

2) TPA：TPA 是完全可信的，它负责验证存储在 CSP 中数据的完整性，在验证过程中不会获得任何具体的数据信息。

3) CSP：CSP 是半可信的，它提供足量的存储空间和检索功能，但同时它是好奇的，并会在多种动机的促使下返回错误的数据，造成机密数据的泄露。

4) Users：群组用户包含一个群管理者和其他用户，管理者具有注册和追踪用户的权利，通过管理者完成注册的用户，可以访问和更新共享数据。

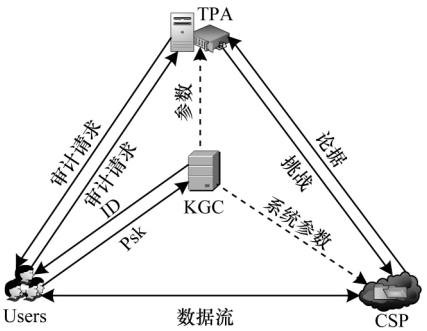

图 1 系统模型

Fig. 1 System model

1.2 安全模型

在本文方案中，CSP 是半可信的，因此，可能会遭受到来自内部或外部的攻击；审计方案的安全模型是由敌手 $\mathcal{A}$ 和挑战者 F 之间的游戏定义的，而该游戏是基于 CDH 困难性问题。挑战游戏有以下阶段：

1) 初始化阶段：挑战者 F 执行初始化阶段，产生系统公共参数 pp 和系统主密钥 msk，然后将公共参数 pp 发送给敌手 $\mathcal{A}$，保留主密钥 msk。

2) 询问阶段：敌手 $\mathcal{A}$ 可以向挑战者 F 做多次询问，包括密钥提取询问和签名生成询问。

(1) 密钥提取询问：挑战者 F 接收到询问的身份信息 ID_i，通过执行运算，将(ID_i，x_i，pk_i)发送给敌手 $\mathcal{A}$。

(2) 签名生成询问：挑战者 F 接收到询问的身份信息 ID_i，数据块 m_i 和数据块 id_i，F 执行签名生成算法，将$\{\sigma_i, C_{i,1}, C_{i,2}\}_{i\in[1,n]}$发送给 $\mathcal{A}$。

3) 输出结果：敌手 $\mathcal{A}$ 输出伪造的签名$\{\sigma_i, C_{i,1}, C_{i,2}\}_{i\in[1,n]}$，若能通过签名验证，则 $\mathcal{A}$ 赢得挑战游戏，CDH 困难问题被攻破。

1.3 设计目标

本文方案需要满足以下性质：

1) 公开审计：公开审计者可以在不检索整篇数据的前提下验证共享数据的完整性。

2) 正确性: 公开验证者可以正确验证共享数据的完整性。

3) 不可伪造性: 只有群组用户可以产生有效、共享数据的原始数据(如签名)。

4) 身份的隐私性: 公开审计者在对共享数据块的审计过程中不能区分签名者的身份信息。

5) 批量审计: 公开审计可以满足多个数据块的一次性验证需求,提高审计效率。

6) 追踪性: 当群管理者发现群组用户有恶意的访问行为时,可以使用追踪权限找到该用户。

2　预备知识

2.1　双线性运算

设 p 为大素数,G_1 是阶为 p 的乘法循环群,g 是 G_1 的生成元。定义一个双线性映射 $e: G_1 \times G_1 = G_2$ 满足以下 3 个性质[20]:

1) 双线性: 对于 $\forall a, b \in \mathbb{Z}_p^*$, $g_1, g_2 \in G_1$,都有 $e(g_1^a, g_2^b) = e(g_1, g_2)^{ab}$。

2) 非退化性: 存在 $\forall P, Q \in G_1$,使 $e(P, Q) \neq 1$。

3) 可计算性: 对于 $\forall P, Q$,可以通过多项式时间算法来计算 $e(P, Q)$。

2.2　困难性问题

本文方案的安全性证明基于以下困难性问题:

定义 1　Computational Diffie-Hellman(CDH)问题。对于 2 个随机数 a、b,已知 (g, g^a, g^b),计算 g^{ab} 的值是困难的,其中 G_1 是 q 阶的循环群,g 是 G_1 的生成元。

定义 2　Decisional Diffie-Hellman(DDH)问题。给定 (g, g^a, g^b, g^c),没有任何的多项式运算可以判定 $g^c = g^{ab}$,其中 $a, b, c \in \mathbb{Z}_p^*$,$G_1$ 是 q 阶的循环群,g 是 G_1 的生成元。

定义 3　Discrete Logarithm(DL)问题。输入 g, $g^a \in G_1$ $(a \in \mathbb{Z}_p)$,输出 a,DL 假设在 G_1 中计算是不可行的,在 G_1 中解决 DL 问题是困难的。

2.3　椭圆曲线上的 ElGamal 加密体制

本文方案使用有效的群签名以获得数据完整性的审计,通过群签名嵌入椭圆曲线上的 ElGamal 加密体制来保护群组用户的私钥。椭圆曲线上的 ElGamal 加密体制安全是基于 DDH 困难性问题而成立的,其加密过程描述如下:

1) 初始化: G_1 是阶为素数 q 的乘法循环群,g 为 G_1 的生成元。随机数 $x \in \mathbb{Z}_q$ 是加密者的私钥,$\mathrm{pk} = g^x \in G_1$ 是其对应的公钥。

2) 加密阶段: 给定消息 $m \in G_1$,加密者首先随机选择暂时密钥 $y \in \mathbb{Z}_q$,计算 $C_1 = g^y \in G_1$, $C_2 = m \cdot \mathrm{pk}^y$,将 (C_1, C_2) 作为密文返回值。

3) 解密阶段: 解密者加密私钥 x,通过计算 C_2 / C_1^x 而得到明文 m。

3　无证书公开审计方案

3.1　单用户情形

本文可证明安全的无证书公开审计方案包括以下 9 个运算阶段。

1) 初始化(Setup)阶段。输入安全参数 λ,由 KGC 生成系统参数,其中,G_1 是阶为素数 q 的循环群,g 是 G_1 的生成元; 存在双线对运算 $e: G_1 \times G_1 \to G_T$,存在 2 个抗碰撞的哈希函数 $h: (0,1)^* \to \mathbb{Z}_p^*$, $H: (0,1)^* \to G_1$。KGC 随机选择 $\alpha \in \mathbb{Z}_q^*$ 作为系统主密钥,计算系统公钥 $P_{\mathrm{pub}} = g^\alpha$,公共参数 $\mathrm{PP} = \{g, G_1, G_T, h, H, P_{\mathrm{pub}}\}$。

2) 密钥生成(KeyGen)阶段。用户 U_i 将自己的 ID_i 发送给 KGC,KGC 根据用户发送的身份信息生成用户的部分私钥 $\mathrm{psk}_{1i} = H(\mathrm{ID}_i)^\alpha$,部分公钥 $\mathrm{pk}_i = H(\mathrm{ID}_i)$,群管理者的部分私钥 $\mathrm{psk}_{1m} = H(\mathrm{ID}_m)^\alpha$,部分公钥 $\mathrm{pk}_m = H(\mathrm{ID}_m)$,KGC 通过安全信道将 psk_{1i} 和 psk_{1m} 分别发送给用户和群管理者; 用户和群管理者在接收 KGC 发送的 psk_{1i} 和 psk_{1m} 后,随机选择 t_i, $t_m \in \mathbb{Z}_q^*$,生成私钥 $\mathrm{psk}_{2i} = H(\mathrm{ID}_i)^{\alpha t_i}$, $\mathrm{psk}_{2m} = H(\mathrm{ID}_m)^{\alpha t_m}$; $T_i = H(\mathrm{ID}_i)^{t_i}$, $T_m = H(\mathrm{ID}_m)^{t_m}$ 分别作为用户和群管理者的公钥; 群管理者随机选择 $d \in \mathbb{Z}_q^*$ 作为自己的专属私钥,计算得到用户的公钥 $D_i = H(\mathrm{ID}_i)^{t_i d}$,群管理者的公钥 $D_m = H(\mathrm{ID}_m)^{t_m d}$ 和 $D = g^d$。

3) 用户注册(UserEnroll)阶段。用户 U_i 将自己的身份信息 ID_i 发送给群管理者,群管理者根据用户 U_i 的 ID_i 和公钥 $\{T_i\}_{i \in [1,n]}$ 建立用户列表 List,完成用户的注册。

4) 签名生成(SignGen)阶段。数据拥有者(Data Owner)将数据存储在 CSP 中,将数据 M 划分成 n 个数据块,即 $M = \{m_1, m_2, \cdots, m_n\}$, $m_l \in \mathbb{Z}_q$; 对每个数据块产生对应的签名,用户 U_i 随机选择 $k_i \in \mathbb{Z}_q^*$,计算得到 $V_i = g^{k_i}$, $h_l = g^{h(\mathrm{id}_l)}$, $l = 1, 2, \cdots, n$,其中 id_l 是数据的标签名,生成对应的签名值: $C_{l,1} = g^{k_i m_l}$, $C_{l,2} = H(\mathrm{ID}_i)^{t_i} D^{k_i m_l} \sigma_l = h_l^{k_i} H(\mathrm{ID}_i)^{\alpha t_i} = g^{h(\mathrm{id}_l) k_i} H(\mathrm{ID}_i)^{\alpha t_i m_l}$。将签名值 $\{\sigma_l, C_{l,1}, C_{l,2}\}_{l \in [1,n]}$ 发送给 CSP 存储进行后续验证。

5) 签名验证(SignVerify)阶段。CSP 接收到用户上传的签名值 $\{\sigma_l, C_{l,1}, C_{l,2}\}_{l \in [1,n]}$ 后对其进行验证。验证式(1)是否成立。若等式成立,则签名有效; 若不成立,则签名无效,需重新生成对应签名。

$$e(\sigma_l, g) = e(h_l, V_i)\, e(T_i^{m_l}, P_{\mathrm{pub}}) \tag{1}$$

6) 挑战生成(ChalGen)阶段。TPA 在接收到用户发送的审计请求后,向 CSP 发送挑战消息。TPA

从 n 个数据块中选取 c 个数据块，其中 $c \in [1,n]$，$j \in [1,c]$，对每个 j 随机选择 $\rho_j \in \mathbb{Z}_q^*$，产生挑战消息 $\text{chal}=\{j,\rho_j\}_{j\in c}$，并将 $\text{chal}=\{j,\rho_j\}_{j\in c}$ 发送给 CSP。

7) 证据生成(ProofGen)阶段。CSP 接收到 TPA 发送的挑战消息 $\text{chal}=\{j,\rho_j\}_{j\in c}$ 后，计算出与之对应的证据：$E=\sum_{j\in c} m_j\rho_j$，$\sigma=\prod_{j\in c}\sigma_j^{\rho_j}$，而后将生成的证据 $\text{proof}=\{E,\sigma\}$ 发送给 TPA 进行验证。

8) 证据验证(ProofVerify)阶段。TPA 接受到 CSP 发送证据 $\text{proof}=\{E,\sigma\}$ 后，对 CSP 生成的证据进行验证。若式(2)成立，则表示数据被 CSP 正确持有；否则，数据可能被篡改或丢失。

$$e(\sigma,g)=e\left(V_i^{\sum_{j\in c}h(\text{id}_j)\rho_j},g\right)e(T_i^E,P_{\text{pub}}) \tag{2}$$

9) 用户追踪(Trace)阶段。如果用户存在恶意的访问行为或上传不合法的数据情况，群管理者可以行使追踪权限，用专属私钥 d 进行追踪，查询到该用户的身份信息并将其从注册用户名单中删除，更新用户列表 List，并将更新后的用户列表发送给 CSP，CSP 删除与之对应存储的签名值，使其不能通过签名验证，从而保证签名的有效性。如果有新的用户想要加入群组，可以通过群管理者进行注册加入。重新对撤销用户生成签名标签的数据块签名，生成新的签名，签名验证过程与之前一致。在群管理者行使追踪权限时，用专属私钥 d 进行追踪，通过下式进行验证：

$$T_i=\frac{C_{l,2}}{C_{l,1}^d}=\frac{H(\text{ID}_i)^{t_i}\cdot D^{k_i m_l}}{(g^{k_i m_l})^d}=H(\text{ID}_i)^{t_i} \tag{3}$$

3.2 多用户情形

当出现多个用户发送审计请求时，TPA 可以在同一时间完成审计。若给定 m 个用户 U_i，$i=1,2,\cdots,m$，每个用户需要审计的数据块数目为 n_i，对应的数据块标识符为 $\text{id}_{i,j}$。TPA 向 CSP 发送审计挑战为：$\text{chal}=\{j,\rho_{i,j}\}$，$j=1,2,\cdots,n_i$。

CSP 产生对应的证据：$\sigma_i=\prod_{j=1}^{n_i}\sigma_{i,j}^{\rho_{i,j}}$，$\sigma=\prod_{i=1}^{m}\sigma_i$，$E_i=\sum_{j=1}^{n_i}m_{i,j}\rho_{i,j}$，之后将生成的证据 $\text{proof}=\{\sigma,E_1,E_2,\cdots,E_m\}$ 发送给 TPA 进行验证。

TPA 接收到 CSP 发送证据 $\text{proof}=\{\sigma,E_1,E_2,\cdots,E_m\}$ 后，验证下式是否成立：

$$e(\sigma,g)=e\left(\prod_{i=1}^{m}V_i^{\sum_{j=1}^{n_i}h(\text{id}_{i,j})\rho_{i,j}},g\right)e\left(\prod_{i=1}^{m}T_i^{E_i},P_{\text{pub}}\right) \tag{4}$$

4 正确性分析与安全性证明

4.1 正确性分析

定理 1 给定共享数据块和对应的合法签名，TPA 可以检验共享数据的完整性。

证明 本文方案的正确性基于式(1)和式(2)。基于双线性对的性质，式(1)和式(2)的验证过程分别如式(5)和式(6)所示：

$$\begin{aligned}&e(\sigma_l,g)=e(g^{h(\text{id}_i)k_i}H(\text{ID}_i)^{\alpha t_i m_l},g)=\\&e(g^{h(\text{id}_i)k_i},g)e(H(\text{ID}_i)^{\alpha t_i m_l},g)=\\&e(g^{k_i},g^{h(\text{id}_i)})e(H(\text{ID}_i)^{t_i m_l},g^{\alpha})=\\&e(h_i,V_i)e(T_i^{m_l},P_{\text{pub}})\end{aligned} \tag{5}$$

$$\begin{aligned}&e(\sigma,g)=e\left(\prod_{j\in c}\sigma_j^{\rho_j},g\right)=\\&e\left(\prod_{j\in c}\left[g^{h(\text{id}_j)k_i}H(\text{ID}_i)^{\alpha t_i m_j}\right]^{\rho_j},g\right)=\\&e\left(\prod_{j\in c}g^{h(\text{id}_j)k_i\rho_j},g\right)e\left(\prod_{j\in c}H(\text{ID}_i)^{\alpha t_i m_j\rho_j},g\right)=\\&e(g^{k_i},g^{\sum_{j\in c}h(\text{id}_j)\rho_j})e(T_i^{\sum_{j\in c}m_j\rho_j},g)=\\&e(V_i^{\sum_{j\in c}h(\text{id}_j)\rho_j},g)e(T_i^E,P_{\text{pub}})\end{aligned} \tag{6}$$

若式(1)和式(2)成立，则公开验证者认为共享数据是完整的。上述通过验证式(1)和式(2)的正确性，证明了本文方案的正确性。

对于多用户情形的批审计，式(4)验证过程如式(7)所示：

$$\begin{aligned}&e(\sigma,g)=e\left(\prod_{i=1}^{m}\sigma_i,g\right)=\prod_{i=1}^{m}e\left(\prod_{j=1}^{n_i}\sigma_{i,j}^{\rho_{i,j}},g\right)=\\&\prod_{i=1}^{m}\left[e\left(\prod_{j=1}^{n_i}g^{h(\text{id}_{i,j})k_i\rho_{i,j}},g\right)e\left(\prod_{j=1}^{n_i}H(\text{ID}_i)^{\alpha t_i m_{i,j}\rho_{i,j}},g\right)\right]=\\&\prod_{i=1}^{m}\left[e\left(\prod_{j=1}^{n_i}V_i^{h(\text{id}_{i,j})\rho_{i,j}},g\right)e\left(\prod_{j=1}^{n_i}H(\text{ID}_i)^{t_i m_{i,j}\rho_{i,j}},g^{\alpha}\right)\right]=\\&\prod_{i=1}^{m}\left[e(V_i^{\sum_{j=1}^{n_i}h(\text{id}_{i,j})\rho_{i,j}},g)e(T_i^{\sum_{j=1}^{n_i}m_{i,j}\rho_{i,j}},P_{\text{pub}})\right]=\\&e\left(\prod_{i=1}^{m}V_i^{\sum_{j=1}^{n_i}h(\text{id}_{i,j})\rho_{i,j}},g\right)e\left(\prod_{i=1}^{m}T_i^{E_i},P_{\text{pub}}\right)\end{aligned} \tag{7}$$

4.2 安全性证明

定理 2 CDH 困难性假设在 G_1 中成立，则任何多项式敌手在本文方案下伪造签名在计算上是不可行的。

证明 如文献[5,17]中所述，使用无证书签名方案的标准安全模型应考虑 2 种类型的敌手，分别称为 I 型敌手和 II 型敌手，具有不同的攻击能力。这 2 类敌手的详细定义如下：

1) I 型敌手：该类型的敌手无法询问 KGC 的主密钥，但能够用其选择的值替换任何实体的公钥(敌手具有此能力的原因为无证书签名方案中没有涉及到证书管理)。

2) II 型敌手：该类型的敌手可以询问 KGC 的主密钥，但不能替换任何实体的公钥(敌手的成功询问表明无证书签名方案中存在秘钥托管问题)。

下文将证明，如果 I 型敌手或 II 型敌手能够用审计方案生成一个伪造签名，那么就存在一个能够解决 G_1 中 CDH 困难问题的算法 F，这与 G_1 中的 CDH 问题在计算上不可行的假设相矛盾。

1) 考虑 I 型敌手的情况。敌手 $\mathcal{A}_{\mathrm{I}}$: 基于审计方案的构造，为在 F 算法模拟的安全游戏中生成伪造的签名，$\mathcal{A}_{\mathrm{I}}$ 需要向 F 算法请求 5 种不同运行阶段的查询，包括系统初始化查询、hash-I 查询、部分私钥提取查询、hash-II 查询和签名查询。同时，敌手能够在游戏中进行公钥替换。在这个游戏中，hash-I 可以看作是一个随机预言模式。算法 F 模拟游戏如下:

(1) 初始化查询: 敌手 $\mathcal{A}_{\mathrm{I}}$ 询问系统的初始化阶段。F 算法设置 $P_{\mathrm{pub}}=g^{a}$，输出并将整个系统参数 $\mathrm{PP}=\{g,G_1,G_T,h,H,P_{\mathrm{pub}}\}$ 返回到 $\mathcal{A}_{\mathrm{I}}$。

(2) hash-I 查询: $\mathcal{A}_{\mathrm{I}}$ 询问签名者标识的 hash-I 结果。F 选择一个随机的 $r\in\mathbb{Z}_q$，然后抛硬币。硬币显示 1，则概率为 p，否则为 0。如果掷币结果为 1，则 F 设置 $H(\mathrm{ID}_s)=p^{r}\in G_1$; 如果掷币结果为 0，则 F 设置 $H(\mathrm{ID}_s)=(p^{r})^{b}\in G_1$。最后，$F$ 将 $H(\mathrm{ID}_s)$ 的结果返回给 $\mathcal{A}_{\mathrm{I}}$。由于 G_1 是一个循环群，a 是 $\mathbb{Z}_q$ 的一个随机元素，p^{r} 和 $(p^{r})^{b}$ 都是 G_1 的元素，p^{a} 和 $(p^{r})^{b}$ 在 G_1 中具有相同的分布，因此，根据 F 返回的 $H(\mathrm{ID}_s)$ 的结果，$\mathcal{A}_{\mathrm{I}}$ 无法区分掷币的结果。

(3) 部分私钥提取查询: $\mathcal{A}_{\mathrm{I}}$ 询问签名者身份 ID_s 上的部分私钥。如果前一个 $\mathcal{A}_{\mathrm{I}}$ 查询中相应的掷币结果为 1，F 将部分私钥输出为 $\mathrm{psk}_s=H(\mathrm{ID}_s)^{r}$; 在相应的 hash-I 查询中随机选取 r; 否则，F 将输出 $\perp$。

(4) 公钥替换: 根据敌手 $\mathcal{A}_{\mathrm{I}}$ 的假设，敌手 $\mathcal{A}_{\mathrm{I}}$ 可以替换任何实体的公钥。具体地，$\mathcal{A}_{\mathrm{I}}$ 首先生成一个随机的 $x_s\in\mathbb{Z}_p^{*}$，并将签名者的公钥设置为 $\mathrm{pk}_s=H(\mathrm{ID}_s)^{x_s}$。然后，$\mathcal{A}_{\mathrm{I}}$ 将 $(\mathrm{ID}_m,x_s,\mathrm{pk}_s)$ 发送给到 F，F 将记录这个密钥替换，以便之后的使用。

(5) hash-II 询问: $\mathcal{A}_{\mathrm{I}}$ 请求对签名者身份 ID_s 返回 hash-II 查询结果，该签名者的公共密钥为 pk_s，数据块 m 和数据块 id_i。F 输出 $h_i=g^{h(\mathrm{id}_i)}$，并将结果返回给 $\mathcal{A}_{\mathrm{I}}$。

(6) 签名询问: $\mathcal{A}_{\mathrm{I}}$ 通过提交 hash-II 询问返回的结果 V_i，请求数据块 m 和数据块标识 id_i 上签名者的签名。如果前一个 hash-I 询问中对应的掷币结果为 1，那么 F 输出的签名为 $\sigma_i=h_i^{k_i}H(\mathrm{ID}_i)^{\alpha x_s}$，在相应的 hash-I 查询中随机选取 r; 否则，F 输出 $\perp$。最终，$\mathcal{A}_{\mathrm{I}}$ 输出一个伪造的签名 σ。在此基础上，F 知道该伪造的对应 hash-I 询问结果是 $H(\mathrm{ID}_s)=(p^{b})^{r}$，而伪造的结果是 $\sigma_i=h_i^{k_i}(H(\mathrm{ID}_i)^{ab})^{r}$。显然，$F$ 可以通过计算得出 p^{ab}。这意味着如果 $\mathcal{A}_{\mathrm{I}}$ 成功伪造一个签名，那么 F 就能够解决 G_1 中的 CDH 问题。

2) 考虑 II 型敌手的情况。由算法 F 模拟的安全游戏中生成伪造签名，敌手还需要请求不同类型的询问，包括初始化询问、hash-I 查询、部分私钥提取查询、hash-II 查询和签名查询。与 I 型敌手的游戏不同，F 应该将主密钥返回给 $\mathcal{A}_{\mathrm{II}}$，但是，$\mathcal{A}_{\mathrm{II}}$ 不能执行公钥替换。在此游戏中，hash-II 被视为一个随机预言模式。在给定 p、p^{a} 和 p^{b} 的情况下，算法 F 模拟游戏如下:

(1) 初始化询问: $\mathcal{A}_{\mathrm{II}}$ 对初始化阶段进行询问。F 生成随机数 $\lambda\in\mathbb{Z}_p$ 作为主密钥和公共参数 $\mathrm{PP}=\{g,G_1,G_T,h,H,P_{\mathrm{pub}}\}$。然后，$F$ 将主密钥和公共参数返回给 $\mathcal{A}_{\mathrm{II}}$。

(2) hash-I 查询: $\mathcal{A}_{\mathrm{II}}$ 询问签名者身份 ID 的 hash-I 结果。F 计算 $\mathrm{pk}_s=H(\mathrm{ID}_s)$，并将 pk_s 的结果返回给 $\mathcal{A}_{\mathrm{II}}$。

(3) 部分私钥提取询问: $\mathcal{A}_{\mathrm{II}}$ 询问签名者身份 ID 上的部分私钥。F 将部分私钥计算为 $\mathrm{psk}_s=H(\mathrm{ID}_s)^{r\lambda}$，并将询问结果返回给 $\mathcal{A}_{\mathrm{II}}$。作为第 II 类敌手的定义，$\mathcal{A}_{\mathrm{II}}$ 不能执行公钥替换。F 将 pk 设置为签名者的公钥。

(4) hash-II 询问: $\mathcal{A}_{\mathrm{II}}$ 对签名者身份、签名者公钥、数据块 m 和数据块标识进行 hash-II 询问的结果。F 生成一个随机的 $r\in\mathbb{Z}_p$，并进行掷币游戏。硬币显示 1，概率为 p_c，否则为 0。如果掷币结果显示 1，F 设置 $h_i=g^{h(\mathrm{id}_i)}=p^{r}$; 如果掷币结果为 0，$F$ 设置 $h_i=g^{h(\mathrm{id}_i)}=(p^{b})^{r}$。最后，$F$ 将 $h_i=g^{h(\mathrm{id}_i)}$ 的结果返回给 $\mathcal{A}_{\mathrm{II}}$。由于 G_1 是循环群，r 是 $\mathbb{Z}_p$ 的一个随机元素，p 和 p^{b} 都是 G_1 的元素，p^{r} 和 $(p^{b})^{r}$ 在 G_1 中具有相同的概率分布，因此 $\mathcal{A}_{\mathrm{II}}$ 不能根据 F 返回的 hash-II 询问结果来区分掷币游戏的结果。

(5) 签名询问: $\mathcal{A}_{\mathrm{II}}$ 询问数据块 m 和数据块标识 id 上的签名。如果上一个 hash-II 询问中的相应掷币结果为 1，则 F 输出的签名为 $\sigma_i=h_i^{k_i}(H(\mathrm{ID}_i)^{a})^{r}$。否则，$F$ 输出 $\perp$。然后，$\mathcal{A}_{\mathrm{II}}$ 输出一个伪造的签名 σ。最后，F 知道该伪造的对应 hash-I 询问结果是 $H(\mathrm{ID}_s)=(p^{b})^{r}$，而伪造的结果是 $\sigma_i=h_i^{k_i}(H(\mathrm{ID}_i)^{ab})^{r}$。显然，$F$ 可以通过计算得出 p^{ab}。这意味着如果 $\mathcal{A}_{\mathrm{II}}$ 成功伪造一个签名，那么 F 就能够解决 G_1 中的 CDH 问题。

综上所述，如果 $\mathcal{A}_{\mathrm{I}}$ 或 $\mathcal{A}_{\mathrm{II}}$ 能够成功伪造签名，那么 F 就能够解决 G_1 中的 CDH 问题，这与 G_1 中 CDH 问题在计算上不可行的假设相矛盾。因此，在本文方案中生成签名的伪造在计算上是不可行的。

定理 3　只要 DL 假设成立，半可信的云存储服务器在本文方案下生成伪造的审计证明在计算上是不可行的。

证明　如定理 2 所证明的，对于不可信云，如果 G_1 上的 CDH 问题是困难的，则在本文提出的机制

下伪造签名是不可行的。在本文中，除了试图在每个数据块上计算伪造签名来生成伪造的审计证明之外，如果不可信的云可以赢得下面的安全游戏（称为 Game1），则它可以伪造损坏的共享数据的审计证明。安全游戏描述如下：

Game1 一个公开的验证者向云服务器发送一个审计挑战 $\text{chal}=\{j,\rho_j\}_{j\in c}$，基于正确的共享数据 M 生成的审计证明应该是 $\text{proof}=\{E,\sigma\}$，它能够通过式（2）的验证。不可信云生成证据为 $\text{proof}'=\{E',\sigma'\}$ 是基于损坏的共享数据 M'，其中 $M\neq M'$ 且为非零的。如果基于损坏的共享数据 M' 的无效证据可以成功地通过验证，则不可信云将获胜；否则，它会失败。

下文将证明如果不可信云可以赢得 Game1，那么就能找到解决 G_1 中的 DL 问题的方案，这与 DL 假设中 G_1 中的 DL 问题在计算上是不可行的相矛盾。假设不可信云可以赢得 Game1，根据式（2）得到：

$$e(\sigma',g)=e\left(V_i^{\sum_{j\in c}h(\text{id}_j)\rho_j},g\right)e(T_i^{E'},P_{\text{pub}})$$

$\text{proof}'=\{E',\sigma'\}$ 是一个正确的审计证明。

定理 4 在公共审计过程中，公开验证者识别共享数据块 c 中所有签名者的身份的概率至多为 $\frac{1}{d^c}$。

证明 对于算法 A，在一个数据块上揭示签名者身份信息的概率是 $1/d$，因为选择的 c 个数据块的签名是独立的，其中 $c\in[1,n]$，公开验证者能区分共享数据中选择的 c 个数据块的所有签名者身份的总概率最多为 $1/d^c$。

在本文方案中，公开验证者知道共享数据中的每个数据块都是由用户各自签名的，因为它需要用户的公共密钥来验证整个共享数据的正确性。然而，它不能区分每个特定数据块上的签名者是谁。因此，公开验证者在揭示私有信息方面不具有特殊的优势，例如在共享数据中对较多的数据块进行签名时，或者特定的数据块经常被不同的组员频繁修改时，公开验证者均不能有效地分辨出签名者具体的身份。根据上文对方案的正确性分析，证明了本文方案支持数据隐私保护。

定理 5 假设一个审计证明 $\text{proof}=\{E,\sigma\}$，只要 DL 假设成立，对于公开验证者来说，揭示本文机制下共享数据中的任何私有数据在计算上是不可行的。

证明 如果生成的证据包含的元素 $E=\sum_{j\in c}m_j\rho_j$，$\sigma=\prod_{j\in c}\left(g^{h(\text{id}_j)k_i}H(\text{ID}_i)^{\alpha t_i}\right)^{\rho_j}$ 是关于数据块 m_j 中所有元素的线性组合，则直接发送到公开验证者，公共验证者可以在收集充足数量的线性组合之后通过求解线性方程来获取数据的内容。为保存私有数据，以随机数作为计算组合元素。公共验证者则用本文提出的方案审计共享数据的完整性。为能够求解线性方程组，公共验证者必须知道 σ、E 的值。然而，给定签名私钥 $\text{psk}_{2i}=H(\text{ID}_i)^{\alpha t_i}$，计算 t_i 与解决 G_1 中的 DL 问题一样困难，这在计算上是不可行的。因此，给出 E 和 σ，公开验证者不能直接获得数据块中所有元素的任何线性组合，并且不能揭示共享数据 M 中的任何私有数据。

5 性能分析

5.1 计算开销

在计算开销方面，将本文方案与文献[18]方案进行分析对比，如表 1 所示。其中，T_{exp} 表示指数运算所需的时间，T_{mul} 表示乘法运算所需的时间，T_{pair} 表示双线性对运算所需的时间，T_{hash} 表示 Hash 运算所需的时间，T_m 表示幂运算所需的时间，d 表示签名的用户数，c 表示挑战的数据块数目，n 表示数据分块数，s 表示每个数据块中包含的子数据块数。从表 1 可以看出，文献[18]方案主要是支持无证书的公共审计方案，但没有签名认证和用户追踪，其计算开销随着数据块的增多而增大，开销远大于本文方案。本文方案支持签名验证、多数据块的批量审计和用户追踪，而在多用户的情形下，审计时间没有太多的增加，且计算开销比文献[18]方案小。

表 1 2 种方案计算开销对比

Table 1 Comparison of computation costs of two schemes

方案	签名生成阶段	证据生成阶段	证据验证阶段
文献[18]方案	$(s+1)T_{\text{exp}}+(s+1)T_{\text{mul}}+T_{\text{hash}}$	$cT_{\text{exp}}+cT_{\text{mul}}$	$3T_{\text{pair}}+(2c+s)T_{\text{exp}}+(2c+s)T_{\text{mul}}+cT_{\text{hash}}$
本文方案	$5T_{\text{exp}}+3T_{\text{mul}}+(n+1)T_{\text{hash}}$	$cT_{\text{mul}}+cT_{\text{exp}}$	$3T_{\text{pair}}+(c+1)T_{\text{exp}}+cT_{\text{hash}}$

5.2 通信开销

为检验云存储中数据的完整性，公开审计者需要将审计挑战发送给云存储，之后云存储将审计证据发送给公开审计者。$|q|$ 是 $\mathbb{Z}_q$ 中元素的大小，$|G_1|$ 是群 G_1 中元素的大小。具体对比结果如表 2 所示。

表 2 2 种方案审计阶段通信开销对比

Table 2 Comparison of communication costs in the audit phase of two schemes

方案	通信开销
文献[18]方案	$(c+1)|G_1|+(c+1)|q|$
本文方案	$c|G_1|+c|q|$

5.3　实验结果

本文采用模拟实验来评估方案的计算开销和通信开销。实验测试环境为: Inter i5-7200U CPU2.50 GHz，8 GB 内存，Windows10 × 64 操作系统，采用 Pairing Based Cryptography (PBC) (http: //crypto. stanford. edu. /pc/) 在 Visual C ++ 6.0 中编译实现。本文方案中主要阶段的运行时间如表 3 所示。

表 3　本文方案各阶段运行时间

Table 3　Runtime of each phase of the proposed scheme　s

方案阶段	运行时间
初始化	0.015 568
密钥提取	0.055 452
签名生成	0.052 124
签名验证	0.048 922
挑战和证据生成	0.007 643
证据验证	0.056 758

基于本文方案与文献 [18] 方案计算开销和通信开销的对比，图 2 和图 3 分别给出了选择不同数量的数据块时签名生成阶段和证据验证阶段运行时间的对比。

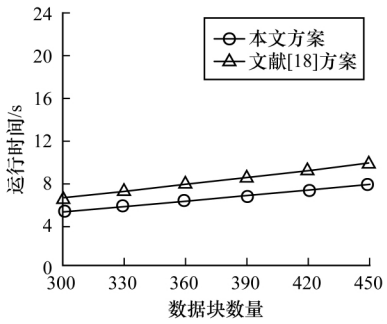

图 2　签名生成阶段运行时间比较

Fig. 2　Comparison of runtime in signature generation phase

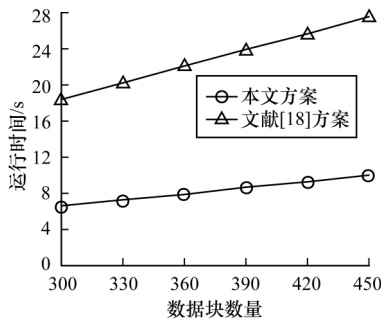

图 3　证据验证阶段运行时间比较

Fig. 3　Comparison of runtime in evidence validation phase

本文方案和文献 [18] 方案都是基于无证书的公开审计方案，因此在证据验证阶段的运行时间相同，但本文方案在签名生成阶段和证据验证阶段的运算时间均少于文献 [18] 方案，从而减少了计算开销，节约了计算成本，提高了效率。

5.4　功能对比

将本文方案与文献 [18-19] 方案的主要功能进行对比，如表 4 所示。其中，√表示支持该功能，× 表示不支持该功能。

表 4　不同方案审计功能对比

Table 4　Comparison of audit functions of different schemes

方案	无证书管理	用户注册	签名验证	用户追踪	批量审计
文献 [18] 方案	√	×	×	×	√
文献 [19] 方案	×	×	×	√	√
本文方案	√	√	√	√	√

从表 4 可以看出，文献 [18-19] 和本文方案同时具有批量审计的功能，但前 2 种方案不支持用户注册和签名验证，且文献 [18] 方案不支持用户追踪，文献 [19] 方案未解决证书管理问题。由此可知，本文方案实现的功能更为全面，不仅解决了传统密码体制中的证书管理问题，而且也实现了批量审计和用户追踪功能。

6　结束语

本文提出一种无证书的共享数据公开审计方案，解决了公钥密码体制中出现的证书管理问题，能够进行多用户多数据块的批量审计，并通过 ELGamal 体制支持群管理者对群组用户身份的追踪。安全性分析和实验结果表明，该方案基于 CDH、DDH 和 DL 困难性问题，实现了签名的不可伪造性和身份的隐私保护性，是安全高效的。下一步将研究具有动态群组用户隐私保护功能的公开批量审计方案。

参考文献

[1] TAN Yuesheng，FAN Wenjie，WANG Jingyu. Privacy-preserving cloud data integrity verification scheme [J]. Journal of Chinese Computer Systems，2017，38(12) : 2736-2740. (in Chinese)
谭跃生，范文婕，王静宇. 一种支持隐私保护的云数据完整性验证方案 [J]. 小型微型计算机系统，2017，38(12) : 2736-2740.

[2] SHACHAM H，WATERS B. Compact proofs of retrievability [J]. Journal of Cryptology，2013，26(3) : 442-483.

[3] HAO Zhuo，ZHONG Sheng，YU Nenghai. A privacy-preserving remote data integrity checking protocol with data dynamics and public verifiability [J]. IEEE Transactions on Knowledge and Data Engineering，2011，23 (9) : 1432-1437.

[4] WANG Cong，WANG Qian，REN Kui，et al. Toward secure and dependable storage services in cloud computing [J]. IEEE Transactions on Services Computing，2012，5(2) : 220-232.

[5] WANG C，CHOW S S M，WANG Q，et al. Privacy-preserving public auditing for secure cloud storage [J]. IEEE Transactions on Computers，2013，62(2) : 362-375.

[6] HUANG Kun,XIAN Ming,FU Shaojing,et al. Securing the cloud storage audit service: defending against frame and collude attacks of third party auditor [J]. IET Communications,2014,8(12):2106-2113.

[7] WANG Zhihao. Cloud storage data integrity verification scheme and its improvement [D]. Chengdu: Southwest Jiaotong University,2018. (in Chinese)
王志豪. 云存储数据完整性验证方案及其改进 [D]. 成都: 西南交通大学,2018.

[8] ZHOU Enguang, LI Zhoujun, GUO Hua, et al. An improved data integrity verification scheme in cloud storage system [J]. Acta Electronica Sinica, 2014, 42(1):150-154. (in Chinese)
周恩光,李舟军,郭华,等. 一个改进的云存储数据完整性验证方案 [J]. 电子学报,2014,42(1):150-154.

[9] ZHA Yaxing,LUO Shoushan,LI Wei,et al. Dynamic group public auditing scheme for shared data on attribute-based threshold signature [J]. Journal of Beijing University of Posts and Telecommunications,2017,40(5):43-49. (in Chinese)
查雅行,罗守山,李伟,等. 基于属性门限签名的动态群组共享数据公开审计方案 [J]. 北京邮电大学学报,2017,40(5):43-49.

[10] HUANG Longxia, ZHANG Gongxuan, FU Anmin. Privacy-preserving public auditing for dynamic group based on hierarchical tree [J]. Journal of Computer Research and Development,2016,53(10):2334-2342. (in Chinese)
黄龙霞,张功萱,付安民. 基于层次树的动态群组隐私保护公开审计方案 [J]. 计算机研究与发展,2016,53(10):2334-2342.

[11] FU Anmin,QIN Ningyuan,SONG Jianye,et al. Privacy-preserving public auditing for multiple managers shared data in the cloud [J]. Journal of Computer Research and Development,2015,52(10):2353-2362. (in Chinese)
付安民,秦宁元,宋建业,等. 云端多管理者群组共享数据中具有隐私保护的公开审计方案 [J]. 计算机研究与发展,2015,52(10):2353-2362.

[12] DOMINGO F J, QIN B, WU Q, et al. Identity-based remote data possession checking in public clouds [J]. IET Information Security,2014,8(2):114-121.

[13] TAN Shuang,JIA Yan. NaEPASC: a novel and efficient public auditing scheme for cloud data [J]. Journal of Zhejiang University-Science C,2014,15(9):794-804.

[14] WANG Huaqun. Identity-based distributed provable data possession in multicloud storage [J]. IEEE Transactions on Services Computing,2015,8(2):328-340.

[15] HUANG Longxia,ZHANG Gongxuan,FU Anmin. Privacy-preserving public auditing for non-manager group [C]//Proceedings of 2017 IEEE International Conference on Communications. Washington D.C.,USA: IEEE Press,2017:1-6.

[16] LIU Hongyu,MU Yi,ZHAO Jining,et al. Identity-based provable data possession revisited: security analysis and generic construction [J]. Computer Standards and Interfaces,2017,54(1):10-19.

[17] ALRIYAMI S S,PATERSON K G. Certificateless public key cryptography [C]//Proceedings of ASIACRYPT'03. Berlin, Germany: Springer,2003:452-473.

[18] WANG Boyang,LI Baochun,LI Hui,et al. Certificateless public auditing for data integrity in the cloud [C]//Proceedings of 2013 IEEE Conference on Communications and Network Security. Washington D.C.,USA: IEEE Press,2013:136-144.

[19] WU L,JING W,ZEADALLY S,et al. Privacy-preserving auditing scheme for shared data in public clouds [J]. Journal of Supercomputing,2018,74(11):6156-6183.

[20] WANG Boyang,LI Baochun,LI Hui. Panda: public auditing for shared data with efficient user revocation in the cloud [J]. IEEE Transactions on Services Computing,2015,8(1):92-106.

编辑 司淼淼

(上接第 142 页)

[14] BANNON L,CYPHER A,GREENSPAN S,et al. Evaluation and analysis of users' activity organization [C]//Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems. New York,USA: ACM Press,1983:54-57.

[15] AHMED A A E,TRAORE I. A new biometric technology based on mouse dynamics [J]. IEEE Transactions on Dependable and Secure Computing,2007,4(3):165-179.

[16] CHEN Guangxin. Proficient in feature engineering [M]. Beijing: People's Posts and Telecom Press,2019. (in Chinese)
陈光欣. 精通特征工程 [M]. 北京: 人民邮电出版社,2019.

[17] SHI Shisong,CHENG Yiming,XIAO Xiaolong. Probability theory and mathematical statistics [M]. Beijing: Higher Education Press,2011. (in Chinese)
茆诗松,程依明,濮晓龙. 概率论与数理统计教程 [M]. 北京: 高等教育出版社,2011.

[18] WU Xizhi,WANG Zhaojun. Nonparametric statistical method [M]. Beijing: Higher Education Press,1996. (in Chinese)
吴喜之,王兆军. 非参数统计方法 [M]. 北京: 高等教育出版社,1996.

[19] WU Xizhi. Statistics: from data to conclusion [M]. Beijing: China Statistics Press,2014. (in Chinese)
吴喜之. 统计学: 从数据到结论 [M]. 北京: 中国统计出版社,2014.

[20] ZHOU Zhihua. Machine learning [M]. Beijing: Tsinghua University Press,2016. (in Chinese)
周志华. 机器学习 [M]. 北京: 清华大学出版社,2016.

[21] ZHOU Tao,HAN Xiaoyu,YAN Xiaoyong,et al. Statistical mechanics on temporal and spatial activities of human [J]. Journal of University of Electronic Science and Technology of China,2013,42(4):481-540. (in Chinese)
周涛,韩筱璞,闫小勇,等. 人类行为时空特性的统计力学 [J]. 电子科技大学学报,2013,42(4):481-540.

编辑 陆燕菲