文章编号:1671-4598(2017)12-0215-04 DOI:10.16526/j.cnki.11—4762/tp.2017.12.056 中图分类号:TP309 文献标识码:A

基于数值分析的匿名大数据访问最优控制算法

周 莉,王伟华,张 敬
(齐齐哈尔大学 理学院,黑龙江 齐齐哈尔 161006)

摘要:为了降低大数据访问对人们生活的影响,减少因数据访问带来的一系列问题,更好地保护用户的隐私,需要对匿名大数据访问进行控制;当前算法是利用Purpose建立匿名大数据访问模型,在原来的K—匿名算法基础上为Purpose匿名数据访问模型构建算法,该算法对公开信息隐私安全涉及较少,对分布式数据隐私的安全保障效果不理想;为此,提出一种基于数值分析的匿名大数据访问最优控制算法;该算法利用MapReduce编程框架对匿名大数据用户的公钥和私钥进行初始化,将计算代理权授权,用户把需要保存的数据以及授权传送给第三方,也就是代理方签名,实现匿名大数据的审计;根据属性群对匿名大数据访问进行控制,系统管理员构建一棵二叉树,通过对称加密算法与属性群路径密钥,加密的群密钥,产生报头消息,根据上述所获结果,管理员对属性群密钥进行生成、更新和分发;实验结果证明,所提算法计算开销、存储开销以及通信开销较低,匿名大数据访问控制的效率高,具有较强的可实践性,为该领域的研究发展提供了支撑。

关键词:数值分析;匿名;大数据访问控制

An Anonymous Large Data Access Control Algorithm Based on Numerical Analysis

Zhou Li, Wang Weihua, Zhang Jing
(College of Science, Qiqihar University, Qiqihar 161006, China)

Abstract: In order to reduce the influence of the big data access to people's life, and brings a series of questions to reduce the data access, better protect the privacy of our users, need for anonymous big data access control. The current algorithm is used in the Purpose to establish anonymous data access model, on the basis of the original K — anonymous algorithm for the Purpose of anonymous data access model construction algorithm, the algorithm and the public information privacy about less, the safety of the distributed data privacy protection effect is not ideal. In this paper, an optimal control algorithm for anonymous large data access based on numerical analysis is proposed. The algorithm using graphs programming framework for big data anonymous user's public key and a private key is initialized, calculates the agency authorization, users of the need to save the data, and authorized to transfer to a third party, namely proxy signature, anonymous big data audit. According to the properties of anonymous data access control, the system administrator to construct a binary tree, through the symmetric encryption algorithm and the path to the key attribute group, group of key encryption, produce the header information, according to the results obtained, the administrator to generate key attribute group, update, and distribution. The experimental results show that the proposed algorithm calculation cost, storage cost and communication overhead is lower, the anonymity of the big data access control with high efficiency, strong practical, provided support to the research development of the field.

Keywords: numerical analysis; anonymity; large data access control

0 引言

随着大数据技术的不断发展,使其成为当今社会非常重要的经济资产[1]。为了更好地利用它们,有偿或者无偿的数据共享是一种趋势。数据访问的控制作为大数据安全分享的重要技术之一,在大数据时代有着不可替代的作用。大数据访问控制是公认的确保数据可以安全共享的手段之一[2]。大数据一般包含了数据的收集、数据的存储、数据的共享以及数据的利用等环节构建的庞大且复杂的网。因为大数据的应用场景不同,导致上述环节与技术也有较大差异[3],因此大数据访问控制要面对的安全问题,也各不相同。在网络飞速发展的背景下,由于大数据具有一些新的特征,当前的大数据访问控制不管是算法还是实施,都需要进行改进和创新,以满足当前应用场景下的安全需求[4]。在这种情况下,如何将大数据访问控制达到最优,成为了当前迫切需要解决的问题[5]。而基于数值分析的匿名大数据访问最优控制算法,可以更好地对保护大数据分享时的隐私,减小因为大数据安全问题,对国家、个人或者组织造成的影响,是解决上述问题的可靠途径[6]。随着信息化以及网络化的发展与应用,匿名大数据访问控制成为了当今社会的热点问题,有关学者对其进行了深刻的研究,同时也出现了很多优秀的成果[7]。

文献[8]提出了一种基于匿名广播加密的匿名大数据

收稿日期:2017-05-15; **修回日期**:2017-05-26。
基金项目:黑龙江省教育厅基本业务专项理工面上项目(135109229);国家科技支撑计划课题(2013BAK12B0803)。
作者简介:周 莉(1976-),女,黑龙江齐齐哈尔人,硕士研究生,副教授,主要从事应用数学方向的研究。

访问控制算法。该算法定义了能够防御自适应敌手侵犯，匿名广播加密的安全模型，在合数阶双线性的环境下通过牛顿插值的多项式，对算法进行构建，在保障用户身份匿名的同时，完成高效的加解密，最后通过基于子群判定假设与合数阶判定双线性中 Diffie－Hellman 假设，在标准的模型下证明了该算法，对自适应敌手有密文机密性与接收者匿名性等特性，但是计算开销大。文献［9］提出了一种基于 MA－ABE 的匿名大数据访问控制算法。该算法利用建立分散授权模型，把属性私钥的产生和中央认证机构进行分离处理，通过数据属性与授权机构分别产生并分发属性的私钥组件，采用基于访问结构树的访问控制策略，高效预防用户间及授权机构间的联合攻击。另外，用户密钥的计算不需要用全球的唯一标识，对匿名用户的跨域数据访问表示支持。该算法耗时较短，但是控制效果不好。文献［10］提出了一种基于 CPABE 的匿名大数据访问控制算法。该算法中数据的拥有者把密文与一个表达性比较强的访问树结构相关联，完成细粒度访问控制。根据用户匿名密钥的发布协议，CPABE 可以实现用户的隐私保护，经过执行该协议，被攻击的授权机构没有办法得到有关用户的全局标识信息，他们就无法利用追踪全局标识收集用户属性。该算法在功能性与安全性方面有一定的优势，但是过程繁琐，耗时长。

针对上述产生的问题，提出一种基于数值分析的匿名大数据访问最优控制算法。实验证明，所提算法可以高效地对匿名大数据访问进行控制，对实际匿名大数据访问控制有重要作用。

1 基于网格虚拟组织的匿名大数据访问控制算法

1.1 数据访问中的双线性映射

双线性映射广泛引用在数据加密和签名等领域中，下面给出了双线性映射定义。

假设 G_1 与 G_2 代表素数 p 的循环，g 代表 G_1 的生成元，那么双线性映射 e 为：

$$G_1 \times G_2 \to G_2 \tag{1}$$

式（1）具有以下性质：

1）双线性特性：对所有的 $u,v \in G_1$ 和 $a,b \in \mathbf{Z}_p$，都有：

$$e(u^a,v^b) = e(u,v)^{ab} \tag{2}$$

其中：u^a 和 v^b 代表双线性映射中的元素，

2）非退化特性：对于生成元 g，始终有 $e(g,g) \neq 1$；

3）可计算特性：对任意的 $u,v \in G_1$，可以在一个多项式的时间内对 $e(u,v)$ 进行计算。

给定 $2l+1$ 个元素构成的向量：

$$\vec{y} = (g,h,g_1,g_2,\cdots,g_l,g_{l+2},\cdots,g_{2l}) \in G_1 \tag{3}$$

其中：$g_i = g^{\alpha^i}$，$\alpha \in \mathbf{Z}_p$ 未知。任意选取随机数 $T \in G_2$，判定 $e(g,h)^{\alpha^{l+1}} = T$ 是否成立，如果对任意一个多项式时间，敌手的优势全部小于可忽略的值，那么判定性 l－BDHE 成立。

1.2 匿名大数据访问安全控制

下面利用一系列游戏来定义匿名大数据访问安全控制，通过敌手 A 与模拟器 B 一起参与进行，假设在游戏中，敌手的优势为可忽略的，那么称分类分级数据属性访问控制算法，在选取标记以及适应性选取密文，攻击下是不能区分的。详细定义如下。

游戏开始前，敌手 A 先给出将要攻击的客体标记 L^*，并访问控制阈值 t^*。模拟器 B 接到敌手 A 选择的挑战客体标记 L^*，模拟器的运行系统构建算法获得系统主密钥 MK 与系统的公开参数 PK，模拟器 B 将 PK 传送给敌手，并将 MK 进行秘密保存。

敌手 A 进行多项式次数，适应性私钥提取以及解密询问，模拟器 B 依据掌握的信息，对私钥提取和解密询问做出响应。私钥提取询问：敌手 A 任意选取用户安全标记 L，满足 $|L \cap L^*| < t^*$，模拟器运行的私钥算法获得相应的用户安全标记 L 私钥 sk，将 sk 回给敌手 A。

解密询问，敌手 A 任选取安全标记 L_i 与对应的消息 M 密文 C_i。模拟器必须输出 C_i 所对应的明文 M。

敌手 A 向模拟器 B，提供等长挑战消息两个，分别是 m_0 和 m_1，模拟器会随机选取 $\beta \in \{0,1\}$，通过初始化时敌手给的 L^*，对 m_β 进行加密操作：

$$CT = Enc(m_\beta,L^*,PK) \tag{4}$$

将上述密文 CT^* 传送给敌手 A。对敌手 A 实施多项式次数的，适应性私钥的提取与加密询问，可是敌手 A 不会交出被挑战的密文 (CT^*,L) 解密查询，或者满足 $|L \cap L^*| \geqslant t^*$ 的 L 私钥提取询问。

最后，敌手 A 给定 β 值的一个猜测 β'，假设 A 给定正确猜想：

$$\beta = \beta' \tag{5}$$

那么称 A 赢了游戏，其中，A 在游戏中的优势可表示为：

$$|\Pr[\beta = \beta'] - 1/2| \tag{6}$$

如果多项式时间中，有任意敌手在游戏中，至多进行 q_K 次的私钥提取查询，以及至多 q_D 次的解密查询，这样的优势可以被忽略，那么就可以判断数据访问控制的很好。

2 基于数值分析的匿名大数据访问最优控制算法

2.1 匿名大数据审计

为了使匿名大数据访问更安全，利用 MapReduce 编程框架对匿名大数据进行审计。

假设，用户选取随机的元素 $x \leftarrow F_d$，以及随机元素 $n \leftarrow R_1$，计算 $j \leftarrow f^x$ 与 $w \leftarrow n^x$，f 代表乘法循环群 G_2 生成元。则分式线性映射 o 为：

$$R_1 \times R_2 \to R_t \tag{7}$$

其中：R_1、R_2、R_t 代表 d 阶乘法循环群。则将私钥定义为：

$$sk = (x) \tag{8}$$

将公钥定义为：

$$dk = (n,f,j,w) \tag{9}$$

将文件划分为 r 块，$H = (c_1,c_2,\cdots,c_r)$，用户对 c_i 进行签名：

$$\sigma_i \leftarrow (N(c_i) \cdot n^{c_i})^x (i = 1,2,\cdots,r) \tag{10}$$

其中：$N(\cdot)$ 代表将二进制字符映射至 R_1 的哈希函数。则签名集合为：

$$\Phi = \{N_i\}_{1 \leqslant i \leqslant n} \tag{11}$$

将 $\{H,\Phi\}$ 发送至云服务器终端，并将本地文件删除。

综上所述，用户请求 TPA（第三方审计）对匿名大数据的完整性进行验证，TPA 选取一个任意数据集合 $I = (s_1,\cdots,s_c)$ 来代表 $[1,r]$，其中 $s_\varepsilon = \pi_{k_{drd}}(\varepsilon)$，$1 \leqslant \varepsilon \leqslant c$，$k_{drd}$ 代表 TPA 为每次匿名大数据审计而随机选取的，π 代表一伪装的随机置换

函数。如果 $s_1 \leqslant \cdots \leqslant s_c$，对每个 $i \in I$，TPA 会选取一个随机值 V_i，TPA 会发送挑战给服务器：$(chal) = \{(i, V_i)\}_{i \in I}$。服务器接到 $chal$ 后，计算 $J = \varphi_{k_{dr\varphi}}(chal) \leftarrow F_d, k_{dr\varphi}$ 代表服务器为匿名大数据每次审计随机选取的，φ 代表一伪随机函数。计算随机掩码：

$$O = (w)^{\varepsilon} \in \boldsymbol{R}_1 \tag{12}$$

样品数据块的线性组合为：$\mu' = \sum_{i=I} j_i c_i$，为更好地保护匿名大数据隐私，计算 $\mu = \mu' + \varepsilon \cdot \theta(R) \in F_d$。与此同时，服务器计算的聚集签名为：$\sigma = \prod_{i \in I} \sigma_i^{j_i} \in \boldsymbol{R}_1$。服务器将 $proof = \{\mu, \sigma, O, N(c_i)_{i \in I}\}$ 传送至 TPA。TPA 接到 $proof$ 后，验证下式是否相等，用来判断匿名大数据的完整性是不是受到了保护。

$$o(\sigma \cdot O^{h(O)}, f) = o\left(\prod_{i \subseteq I} N(c_i)^{j_i} \cdot n^{\mu}, jj\right) \tag{13}$$

如果有 χ 个用户需要对匿名大数据的完整性进行验证，假设用户 δ 有文件 $H_\delta = (c_{\delta+i}, \cdots, c_{\delta+r})$，对于用户 δ，选取一个随机 $x_\delta \in F_d$ 当作私钥，则对应的公钥为：

$$(j_\delta, w_\delta, f_\delta, n_\delta) = (f^{x_\delta}, n^{x_\delta}, f_\delta, n_\delta) \tag{14}$$

其中：j_δ、w_δ、f_δ 和 n_δ 代表私钥中的元素，f^{x_δ}、n^{x_δ}、f_δ 和 n_δ 代表对应的公钥。用户 δ 的匿名数据块 $c_{\delta,i}$ 的签名为：

$$\sigma_{\delta,i} \leftarrow [N(\delta \| c_i \|) \cdot (n_\delta)^{c_{\delta,i}}]^{x_\delta}, i \in [1, r] \tag{15}$$

服务器接收挑战 $chal$ 后，将会任意为每个用户选取 J_δ，并对 J_δ 进行计算：

$$\mu_\delta = \prod_{i \subseteq I} V_i c_{\delta,i} + J_\delta \theta(O_\delta) \in F_d \tag{16}$$

$$\mu_\delta = \prod_{i \subseteq I} V_i c_{\delta,i} + J_\delta \theta(O_\delta) \in F_d$$

$$\sigma = \prod_{\delta=1}^{\chi} \left(\prod_{i=s_1}^{s_c} (\sigma_{\delta,i}^{j_i}), \right) O_\delta = (w_\delta)^{J_\delta} =$$

$$(n_\delta^{x_\delta})^{J_\delta} \circ proof = \{\sigma, \{\mu_\delta\}_{1 \leqslant \delta \leqslant \chi}, \{O_\delta\}_{1 \leqslant \delta \leqslant O}, N(\delta \| c_i \|)_{i \in I}\} \tag{17}$$

将上式当作证据还给 TPA，TPA 收到证据后，依据 $proof$ 与公钥验证式，一次性对 χ 个用户的匿名数据，完整性进行判断。下式为公钥验证式：

$$o\left(\sigma \cdot \prod_{\delta=1}^{\chi} O_\delta^{\theta(O_\delta)}, f\right) = \prod_{\delta=1}^{\theta} o\left(\prod_{i=s_1}^{s_c} [N(\delta) \| c_i \|]^{j_i} \cdot (n_\delta)^{\mu_\delta}, j_\delta\right) \tag{18}$$

2.2　匿名大数据访问控制

以 2.1 中的审计结果为基础，利用属性群对匿名大数据访问进行控制。

假设数据服务的管理员接到密文 (C_τ, CT) 之后，对 CT 进行下列操作：

对每个属性 $\rho \in \eta$ 所对应属性群 U_{λ_ρ}，任意选取 $Y_{\lambda_\rho} \in \vartheta_v^*$ 当作 U_{λ_ρ} 的群密钥，对 CT 进行计算：

$$CT = (T, \widetilde{C} = \kappa_\tau \partial(\omega, \omega)^{\omega^s}, C = \omega^s, \forall \rho \in \eta) \tag{19}$$

管理员构建一棵二叉树，也可称为 KEK 树。在该树中找到可以覆盖 U_{λ_ρ} 中，所有用户中的最小子树，将这些子树根节点所对应的任意数定义为，属性群 U_{λ_ρ} 路径密钥，表示为 $KEK(U_{\lambda_\rho})$。假设属性群 $U_{\lambda_1} = \{\zeta_1, \zeta_3, \zeta_4\}$，那么只有 U_{λ_1} 的用户才有权知道 $KEK(U_{\lambda_1})$。

通过对称加密算法 E 与属性群路径的密钥加密的群密钥，产生报头消息：

$$Hdr = (\forall \rho \in \eta : \{E_K(K_{\lambda_\rho})\}_{K \in KEK(U_{\lambda_\rho})}) \tag{20}$$

管理者在数据服务器上保存 (Hdr, C_τ, CT)。

用户访问大数据文件 τ 过程如下：解密报头消息 Hdr 获得群密钥，解密密钥密文 CT 获得对称的密钥 k_τ，解密根据数据密文 C_τ 获得数据文献 τ。详细过程如下：

如果用户 u_t 向云服务提供商发起匿名大数据文件 τ 访问请求，管理者向 u_t 将服务器上的 (Hdr, C_τ, CT) 返还。

u_t 解密 Hdr 获得所属的属性群，对应的群密钥，假设用户 u_t 有一合法属性 λ_ψ，那么 u_t 通过 $KEK \in KEK(U_{\lambda_\psi}) \cap PK_t$ 解密 Hdr 获得属性群 U_{λ_ψ} 所对应的群密钥 K_{λ_ψ}，由此 u_t 可更新私钥：

$$SK_t = (\xi = \omega^{(\alpha+r)/\lambda}, \forall \lambda_\psi \in \vee) \tag{21}$$

u_t 利用新的私钥解密密文 CT，对 T 中的叶节点 ρ，定义递归函数：

$$DecryptNode(CT, sk, \rho) = \left\{\frac{\partial(D_\rho, C_\rho)}{\partial(D'_\rho, C'_\rho)}\right\} = \partial(\rho, \rho) \tag{22}$$

对于非节点 $\bar{\omega}$：假设 $\bar{\omega}$ 的子节点集为 $\{z_j\}$，对每个子节点 $\{z_j\}$ 所对应的 $F_{z_j} = DecryptNode(CT', sk, z_j) = \partial(\omega, \omega)$，选择 k_ω 个子节点中的 F_{z_j} 当作 Lagrange 插值多项式中的插值节点，并计算 $F_\omega = \partial(\omega, \omega)$，对于根节点 ζ，设定 $R = DecryptNode(CT', sk, \zeta) = \partial(\omega, \omega)$，假设 u_t 属性集 Λ 满足访问结构树 Ψ，那么对称密钥 $k_\tau = \widetilde{C} / (\partial(C, D)/A)$。$u_t$ 通过解密 C_τ 获得数据明文 τ。

3　仿真实验结果与分析

为了证明基于数值分析的匿名大数据访问最优控制算法的有效性，需要进行一次实验，在 linux 的环境下搭建匿名大数据访问控制实验仿真平台。实验数据取自于中软卓越大数据分析公司，利用本文所提算法将该实验在 10 台 PC 机，CPU 为 4.00 GHz，RAM 为 2048 MB 的硬件下进行实验，由此观察本文所提算法的整体有效性。表 1 是不同算法明文加密时间（s）对比。

表 1　不同算法明文加密时间对比

明文大小/KB	文献[8]算法所用时间/s	文献[10]算法所用时间/s	本文算法所用时间/s
2	1.0	1.5	0.4
4	1.6	2.6	0.9
6	2.2	3.5	1.2
8	2.9	4.8	1.7
10	3.5	5.2	2.1

分析表 1 可知，文献［8］所提算法在合数阶双线性的环境下通过牛顿插值的多项式，对访问控制算法进行构建，增加了明文加密时间，不适用于大规模明文加密。文献［10］所提算法利用用户匿名密钥的发布协议，通过 CPABE 实现用户的隐私保护，这一步骤本身就存在延时，所以导致明文的加密时间较长。而本文所提算法利用对称加密算法加密大规模匿名数据，通过属性加密算法对对称密钥进行加密，减少了明文加密时间。证明了本文所提算法具有可行性。表 2 是在数量不等的

大数据属性分类数目（个）下，不同算法密文所占存储空间（KB）对比。

表 2 不同算法密文所占存储空间对比

数据属性分类数目/个	文献[9]算法密文所占存储空间/KB	文献[10]算法密文所占存储空间/KB	本文算法密文所占存储空间/KB
5	0.45	0.54	0.27
10	0.47	0.58	0.31
15	0.52	0.62	0.34
20	0.57	0.84	0.39
25	0.60	0.91	0.43
30	0.64	0.99	0.47
35	0.71	1.20	0.51

由表 2 可知，不同数量的数据属性下，文献［9］所提算法下密文所占存储空间，与本文所提算法下密文所占存储空间相比略大，文献［9］所提算法把属性私钥的产生和中央认证机构进行分离处理，通过数据属性与授权机构分别产生并分发属性的私钥组件，增加了密文所占的存储空间。在文献［10］所提算法中，数据的拥有者把密文与一个表达性比较强的访问树结构相关联，该访问树结构所占内存偏大，导致文献［10］所提算法，密文所占存储空间较大。本文所提算法把重加密任务交给了云服务提供商，不仅降低了匿名数据拥有者计算代价，而且减少了密文所占存储空间。进一步证明了本文所提算法具有实用性和适用性，是一种值得借鉴的匿名大数据访问控制算法。图 1 是不同算法数据审计时间（s）对比。

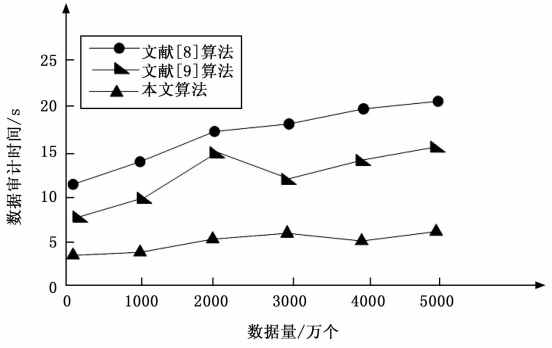

图 1 不同算法数据审计时间对比

由图 1 可知，在数据审计所用时间方面，本文所提算法明显优于文献所提算法。本文所提算法利用 MapReduce 编程框架对匿名大数据进行审计，不仅将用户的开销降到了最低，近乎常量，而且还对大数据服务的运算效率与服务性能进行了考虑和分析，减少了数据审计时间，为匿名大数据访问控制提供了基础。说明了本文所提算法具有较强的实践性。图 2 是不同算法匿名大数据访问控制效果对比。下图是数据访问控制效率（%）的对比，下式为控制效率（%）计算公式：

$$\varphi = \frac{E_K U_{\lambda_i} \widetilde{C}}{\vartheta^* \partial(\omega,\omega)} \times 100\% \tag{23}$$

分析图 2，文献所提的三种算法匿名大数据访问控制效果一般。文献［8］所提算法和文献［9］所提算法在访问次数小于 400 次时，匿名大数据访问控制效率曲线平缓，但当匿名大数据访问次数大于 400 次时，匿名大数据访问控制效率下降趋势明显，文献［10］所提算法在上图中，有两个阶段数据访问控制效率有降低趋势，说明了文献所提算法通用性较差。本文所提算法首先对匿名大数据进行了审计，然后进行访问控制，提高了数据访问控制效率，为该领域发展提供了基础。

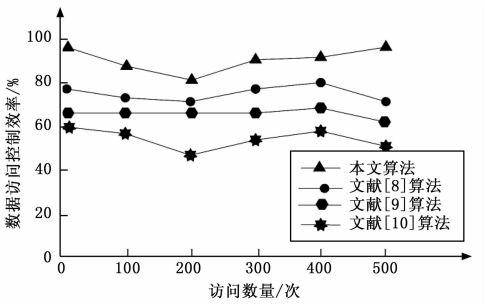

图 2 不同算法匿名大数据访问控制效率对比

实验证明，本文所提算法可以快速、安全地对匿名大数据访问进行控制，减少了数据丢失率，最大程度地保障了用户隐私，大大降低了数据被恶意篡改的可能性，具有抗合谋攻击性、后向保密性和数据机密性以及前向保密性。能够友好地和对象存储管理机制相结合，有较高的实际意义。

4 结束语

由于数据的安全性以及管理性都面临着全新的挑战，采用当前算法对匿名大数据访问进行控制时，存在动态复杂性高，开放性与资源高度集中带来的不安全性得不到解决的问题。提出一种基于数值分析的匿名大数据访问最优控制算法。并通过实验证明，所提算法可以高精度，高效率地对匿名大数据访问进行控制，是一种可借鉴的匿名大数据访问控制算法，是一种安全、灵活、可靠地算法，对保护用户数据的机密性有很重要的作用。

参考文献：

[1] 朱彦杰. 基于最优化网格分配的资源数据库访问控制［J］. 科技通报，2015，31（8）：216－218.
[2] 刘莉苹，李国庆. 基于属性的空间数据访问控制研究［J］. 计算机工程与设计，2014，35（3）：803－808.
[3] 田 野，彭彦彬，杨玉丽，等. 无线体域网中基于属性加密的数据访问控制方案［J］. 计算机应用研究，2015，32（7）：2163－2167.
[4] 王莉莉，周 雄. 基于主题模型的数据库访问准确性优化研究［J］. 计算机仿真，2016，33（10）：450－453.
[5] 周明快. 基于 CP－ABE 的云计算改进属性加密安全访问控制策略设计［J］. 计算机测量与控制，2015，23（1）：297－299.
[6] 胡 雷，杨剑锋，郭成城，等. 多信道无线 Mesh 控制网中双路径并发性能研究［J］. 科学技术与工程，2015，15（4）：150－154.
[7] 惠 榛，李 昊，张 敏，等. 面向医疗大数据的风险自适应的访问控制模型［J］. 通信学报，2015，36（12）：190－199.
[8] 何文婷，刘 健，袁庆升，等. 支持 Hadoop 大数据访问的 pNFS 框架研究与实现［J］. 计算机应用研究，2016，33（11）：3340－3344.
[9] 阎 栋. 基于 ADO 数据库访问技术的会计电算化系统设计与实现［J］. 电子设计工程，2016，24（23）：159－161.
[10] 曹珍富，董晓蕾，周 俊，等. 大数据安全与隐私保护研究进展［J］. 计算机研究与发展，2016，53（10）：2137－2151.