一种新型的安全云存储模型

王玉玺[1]，张串绒[1]，张柄虹[1]，张琳琳[2]，许长鹏[3]
(1.空军工程大学信息与导航学院,陕西西安,710077;2.空军大连通信士官学校,辽宁大连,116000;
3.71320 部队,河南开封,475000)

摘要　针对云存储中的数据安全访问控制问题,设计了一个以云服务器为中心节点,支持不同权限多用户访问的安全高效的云存储模型。该模型利用椭圆曲线公钥密码设计具有语义安全的可搜索加密算法,并结合关键字相关度定义和保序加密算法提高数据检索准确性,降低系统通信负担;通过代理加密思想、用户访问控制列表和用户——数据访问控制矩阵实现不同权限用户的访问控制,安全灵活解决动态用户密钥管理问题。最后通过对模型安全性和模型效率两个方面对比分析,对模型可行性给出了证明。

关键词　云存储;椭圆曲线密码;保序加密;密钥管理;可搜索加密
DOI　10.3969/j.issn.1009-3516.2013.05.017
中图分类号　TN918.1　　**文献标志码**　A　　**文章编号**　1009-3516(2013)05-0071-05

A New Secure Model of Cloud Storage

WANG Yu-xi[1],ZHANG Chuan-rong[1],ZHANG Bing-hong[1],ZHANG Lin-lin[2],XU Chang-PENG[3]
(1.Information and Navigation College, Air Force Engineering University, Xi′an 710077, China;
2. Air Force Communications School of Warrant Officers, Dalian 116000,Shenyang,China;
3. Unit 71320,Kaifeng 475000,Henan,China)

Abstract: Aimed at the problem of the secure access to the data in cloud storage, an efficient and secure model of cloud storage which supports the access of multi-user in hierarchy and adopts the cloud server as the central point is proposed. This model is used to design a semantic searchable encryption with the use of elliptic curve cryptosystem, and the use of the model improves the accuracy of the result by the keyword's degree of correlation and the order preserving encryption. Through the combination of proxy encryption, user access control list and user-data access control matrix, this model can satisfy the access of multi-user in hierarchy and efficiently solve the management of the dynamic user′s private keys. In the end through the analysis and comparison of the security and efficiency, the results show that his new model is feasible.

Key words: cloud storage; elliptic curve cryptosystem; order preserving encryption; key management; searchable encryption

云存储作为云计算一种重要的服务模式,可以向用户提供按需的、实时的、可扩展的存储服务,不仅极大方便了用户对数据的应用需求,而且为用户节省了大量的建设成本和管理费用。在实际的应用中,数据的机密性、可用性成为云存储模型必须解决的主要问题,支持不同访问权限的多用户安全云存储模型成为人们研究的重点。

为了保证云存储数据的机密性,目前最为有效的解决方案为基于加密方法的数据存储方案为了实现数据在密文安全条件下被合法用户访问利用,

收稿日期:2013-04-24
基金项目:国家自然科学基金资助项目(61272486)
作者简介:王玉玺(1989－),男,山东寿光人,硕士生,主要从事密码学与网络安全研究.E-mail:WYX10013@163.com

Song 等[1]首次提出了可检索的数据加密方法,在该方法中云服务商通过布尔函数验证密文中特定位置的关键字实现特定位置的密文检索;Masayuki 等[2]在文献[1]的基础上通过定义不可识别性,利用同态函数设计语义安全的可搜索加密,解决了[1]中存在的通过对询问内容在密文中的分布对数据进行统计分析的安全隐患;Boneh 等[3]首次提出了基于关键词的公钥可搜索加密算法;Golle 等[4-6]提出了多关键词连接的可搜索加密。现有的可搜索加密方案大都仅满足单一用户与云服务器之间的交互,为实现数据的共享,Vimercati 等[8-9]提出了不同的云计算环境中多用户访问控制模型,为了保证数据机密性与可用性,模型要求数据拥有者作为中心节点对用户访问请求进行受理,因此要求数据拥有者必须时时在线,当用户访问量增大时,资源有限的数据拥有者将成为整个模型的瓶颈。Yang[10]提出了以云服务器为中心节点、支持多用户访问的可搜索加密方案,但该方案中所有用户具有相同的访问权限和解密密钥,对于用户的动态变化管理复杂。为实现不同权限的多用户数据访问,现有的主要解决方法为针对不同权限用户发放不同等级的密钥,使用户按照数据拥有者对用户和数据的逻辑划分进行访问,因此基于不同访问权限的密钥管理成为研究重点。现有的密钥管理主要有文献[11～12]等提出的密钥分发策略和文献[13～15]等提出的密钥推导策略,2 种策略主要集中于动态用户的密钥管理和数据的再加密。

通过上述现有的工作研究发现:目前没有一个以云服务器为中心节点,能够支持密文检索的不同权限的多用户安全访问控制模型。针对该问题,本文设计了一个新型的安全云存储模型。

1 安全理论基础

椭圆曲线公钥密码[16]是建立在求解 ECDLP 的困难性上的,除了少数几类曲线外,到目前为止对于 ECDLP 不存在亚指数算法,因此 ECC 算法能够以较短的密钥得到较高的安全级别。另外由于椭圆曲线上所有的点形成的集合在数学形成群的关系,两个循环群之间存在相对应的双线性映射关系,因此双线性配对函数能够较好地应用于椭圆曲线上。双线性对(Bilinear Pairing)具体定义如下:

定义 1 令 G_1 为由 P 生成的循环加法群,阶为 q,G_2 为具有相同阶 q 的循环乘法群,a,b 为 Z_q^* 中的元素。假设 G_1 和 G_2 2 个群中的离散对数问题是困难的,双线性对是指满足下列性质的一个映射 e:$G_1 \times G_1 \rightarrow G_2$:

1)双线性:$\forall\ P,Q\in G_1$,$\forall\ a,b\in Z_q^*$,$e(aP,bQ)=e(P,Q)^{ab}$;

2)非退化性:如果 g 为群 G_1 的生成元,则 $e(g,g)$为群 G_2 的生成元;

3)可计算性:对于所有 $P,Q\in G_1$,总存在有效的计算方法计算 $e(P,Q)$。

双线性对目前能够通过椭圆曲线或超椭圆曲线中的 Weil 对、Tate 对变形得到[17]。

本文提出的方案其安全性还基于计算性 Diffie-Hellman 问题和双线性 Diffie-Hellman 问题的难解性,具体定义如下:

定义 2 计算性 Diffie-Hellman 问题:给定(P,aP,bP),计算 $abP\in G_1$,其中 $a,b\in Z_q^*$ 是未知整数。

定义 3 双线性 Diffie-Hellman 问题:给定(P,aP,bP,cP),计算 $e(P,P)^{abc}\in G_2$,这里 $a,b,c\in Z_q^*$ 是未知的整数。

2 新型的云存储访问控制模型框架

新型的云存储访问控制模型包含 3 个实体:数据拥有者(Data Owner, DO)、用户(User)、云服务器(Cloud Server, CS)。模型框架见图 1。

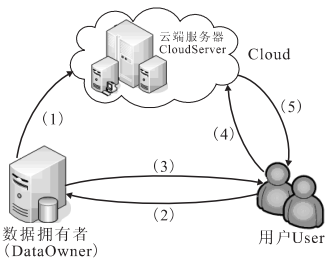

图 1 新型云存储访问控制模型框架

Fig.1 Framework for a new model of cloud storage access control

以下介绍数据拥有者、用户和云服务器之间的交互协议。

1)数据拥有者对每个数据块 D_i,提取关键字 k_i 并根据统计公式计算每个关键字相关度 R_i,对明文数据、关键字和相关度分别采用不同加密算法加密;对合法注册用户进行密钥分发,按照服务内容和用户权限的逻辑划分生成用户访问控制列表(User Access Control List, UACL)和用户——数据访问控制矩阵(User-Data Access Control Matrix, UDACM);将加密结果及 UACL、UDACM 发送给 CS。

2)用户向 DO 注册,User 将自身身份标识加密后发送至 DO。

3)数据拥有者验证用户身份,分发密钥,按照逻

辑划分确定其访问权限。

4)用户向 CS 发起查询请求,希望 CS 返回准确的数据查询结果。

5)服务器按照 UACL、UDACM,利用检索算法查找 User 访问权限内匹配数据,并进行关键字相关度比较,返回相关度最大的 k 项匹配数据。

在构建的模型中有 2 点需要说明:①假设 CS 具有足够的存储空间及理想的计算性能,能够及时处理并响应 User 的访问请求。而且 CS 具有半诚实模型(Curious-But-Honest)的特点,即 CS 准确完成既定的协议并保证不与恶意用户串通攻击存储数据,但是会记录下中间所有结果并对存储的加密信息进行分析,以推导出额外的有价值信息;②为了简单起见,假设每个数据块只提取一个关键字。

3　模型构建中的关键问题

3.1　用户密钥管理

为了实现灵活、细粒度的访问控制机制,每个数据都采用不同的加密密钥 k_{S_i},而且为了降低计算复杂度,利用对称加密算法对明文数据进行加密。为实现不同用户具有不同的访问权限,并能够灵活解决动态用户的密钥管理问题,避免用户动态变化而对数据重新加密,本文采用 UACL 和 UDACM 控制方法,将用户访问权限的逻辑控制和密钥管理由 DO 和 CS 协作实现,大大减轻了系统动态用户密钥管理的复杂度。模型密钥管理算法具体如下:

STEP 1:模型初始化,选取素数域 $GF(p)$上满足安全要求的椭圆曲线 C,及阶为 q 的群 G_1 和 G_2,G_1 的生成元为 P,其中 $q\geqslant 163$ bits。

STEP 2:DO 选取私钥 e_o 作为主密钥,满足 $\gcd(e_o,q)=1$。对于注册用户 U_i,DO 选取参数 e_i,d_i 满足 $e_i d_i\equiv 1 \bmod q$,计算 e_i' 使得 $e_o e_i'\equiv e_i \bmod q$。因为 $\gcd(e_o,q)=1$,所以方程 $e_o x\equiv 1 \bmod q$ 有解,进而容易求得 $e_i'\equiv e_i x \bmod q$。同理选取整数 d_i',求得 d_{U_i} 满足 $d_i' d_{U_i}\equiv d_i \bmod q$,用户 U_i 私有密钥为 d_{U_i}。

STEP 3:DO 对所有合法用户生成 UACL,具体见表 1。

表 1　用户访问控制列表

Tab.1　User access control list

用户证书	代理加密密钥
C_1	$d_{U_1}'\equiv e_1' d_1' \bmod q$
C_2	$d_{U_2}'\equiv e_2' d_2' \bmod q$
…	…
C_n	$d_{U_n}'\equiv e_n' d_n' \bmod q$

STEP 4:DO 根据服务策略,按照数据密级和用户权限划分,生成 UDACM 用以表示用户与数据之间的逻辑关系,见表 2。

表 2　用户-数据访问控制矩阵

Tab.2　User-data access control matrix

	D_1	D_2	D_3	…	D_{n-1}	D_n
C_1	1	0	1	…	1	0
C_2	0	0	1	…	1	1
C_3	1	0	1	…	0	1
…	…	…	…		…	…
C_{n-1}	1	1	0	…	0	0
C_n	0	0	1	…	0	1

表 2 中“1”表示用户可以访问相应的数据,“0”表示用户不能访问相应的数据。

3.2　关键字加密检索算法

为充分利用云计算资源,文中设计了以云服务器为中心节点的基于椭圆曲线公钥密码的可搜索加密算法,该算法将数据检索工作交由 CS 完成,大大减轻了 DO 的计算负担。关键字加密检索算法分为 4 个部分:

1)关键字的提取及相关度的计算:DO 为每一个数据进行关键字提取,通过对明文数据 D_i 进行统计分析,提取明文数据的中心词作为关键字,为方便分析假设每个数据 D_i 仅提取一个关键字;利用文献[7]中的相关度静态计算方法,计算:

$$R_i(k_i,D_i)=\frac{1}{|D_i|}(1+\ln f) \tag{1}$$

式中:f 为关键字 k_i 在数据 D_i 中的频率;$|D_i|$表示明文数据长度。

2)索引生成算法 E_{KWEnc}:DO选取函数 $F:\{0,1\}^* \rightarrow G_1$,将原始明文对应到群 G_1 相应离散点上,该过程是可逆的;随机选取整数 r_i 对每个 k_i 计算 $r_i F(k_i)$,并利用主密钥 e_o 对生成元 P 加密生成 $e_o r_i P$,密文对$(r_i F(k_i), r_i e_o P)$即为密文索引,利用保序加密算法[21]计算 $E_{\mathrm{ope}}(R_i,k_{\mathrm{ope}})$作为密文索引属性。

3)询问门限生成算法 E_{query}:用户 U_i 利用密钥 d_{U_i}将所要搜索的关键词 k_i' 进行加密,计算得到 $d_{U_i} F(k_i')$,随机选择整数 l 并计算 $l d_{U_i} F(k_i')$,同时对生成元 P 计算 lP,密文对$(l d_{U_i} F(k_i'), lP)$即为密文查询门限。

4)关键字检索:CS 收到 M_{UC} 之后,根据 UACL 判定用户 U_i 是否为合法用户,若为合法用户则对询问门限进行代理加密 $E_{\mathrm{proxy}}=l d_{U_i}' d_{U_i} F(k_i')=l e_i' d_i' d_{U_i} F(k_i')$;根据 UDACM 确定用户可访问数据范围,在可访问数据范围内 CS 利用双线性对计算并验证等式 $e(l e_i' d_i' d_{U_i} F(k_i'), r_i e_o P)=e(r_i F(k_i), lP)$是否成立。由双线性对性质可知,若 $k_i'=k_i$ 则

有:$e(le_i' d_i' d_{U_i} F(k_i'), r_i e_o P) = e(F(k_i'), P)^{le_i' d_i' d_{U_i} e_o r_i} = e(F(k_i'), P)^{lr_i} = e(r_i F(k_i), lP)$。对所有匹配数据加密后的关键字相关度密文 $E_{ope}(R_i, k_{ope})$进行比较,得到相关度最大的前 k 项作为检索结果。

3.3 数据加密及动态操作

文中设计的模型中,每个明文数据都采用不同的加密密钥 k_{S_i},利用安全高效的对称加密算法对数据 D_i 进行加密,DO 利用主密钥 e_o 计算 $E_{KEnc}(k_{S_i}, e_o) = e_o F(k_{S_i})$,密文对$(E_{KEnc}(k_{S_n}, e_o), E(D_n, k_{S_n}))$即为加密结果。数据的动态操作一般包括数据的增加、删除和更新,在现有的不同权限的多用户密钥管理方案中,数据的动态变化会造成相关用户密钥的更新,工作量大。本文设计的模型中,对于数据的更新数据拥有者仅需利用新的对称密钥 k_{S_i} 将更新后的数据加密,同时对更新后的对称密钥利用 E_{KEnc} 重新加密并代替原有密文。对于数据访问权限发生变化的情况,数据拥有者仅需更新相应的 UDACM,通过半诚实的 CS 代理加密机制[8]控制下实现用户访问权限的变更。

3.4 动态用户密钥管理

在实际应用中,用户因为自身需求的变化其访问权限也会发生实时的动态变化。用户动态变化一般包括新用户的注册、用户的注销和用户访问权限的变化。现有的不同权限的多用户访问控制模型中,为安全解决用户动态变化带来的影响需要对其他用户重新分配密钥,同时对受影响数据重新加密,模型动态用户管理复杂。本文设计的新模型中,每个用户按照密钥生成算法获得互不相关的不同的私有密钥,用户之间的逻辑关系及用户自身的访问权限由半诚实的 CS 根据 DO 生成的 UACL、UDACM 进行管理。新用户注册,DO 将生成密钥分发给新的用户之后,将 UACL 和 UDACM 增加一条相应记录;原有用户注销,DO 只需将 UACL 和 UDACM 中对应记录删除,失去 CS 代理加密服务后的 User 无法进行正常的数据访问;用户访问权限变更,DO 根据变更后用户权限更新 UDACM 中相应记录。

4 模型可行性分析

4.1 模型安全性分析

模型的安全性包括 2 个部分:数据安全性、密钥安全性。本文提出的新的模型其安全理论基础为椭圆曲线上的离散对数,而且在模型 3 个主体交互过程中,前提假设为 CS 具有半诚实模型特点,不会与恶意用户串谋攻击存储在云服务器中的加密数据,同时在整个访问过程中数据拥有者主密钥 e_o、用户密钥 d_{U_i} 作为私钥得到保护。DO 利用不同的对称密钥对每个明文数据进行加密,使数据对 CS 具有不可见性,利用代理加密方法,用户仅靠自身私钥无法正确检索解密明文数据,保证数据机密性;在密文索引和询问门限生成算法中,模型引入随机数设计具有语义安全的随机加密算法 E_{KWEnc} 和 E_{query} 使索引和询问门限具有不可分辨性,避免了 CS 通过对关键词密文索引及询问门限的分布特点进行统计分析而造成的明文信息泄露;为提高检索结果准确性,在检索算法中引入关键字相关度作为密文索引属性,利用保序加密算法在加密条件下实现相关度大小比较,确保明文信息机密性;计算性 Diffie-Hellman 问题防止了云服务商通过自身代理加密密钥破解关键词索引。在该模型中,每个用户都具有各不相关的独立私有密钥,能够较好的解决文献[10]中存在的恶意用户冒名访问和中间人攻击。

4.2 模型效率分析

为证明新的模型具有高效的密文检索效率,本文通过理论分析,将新模型检索算法和文献[10]所提出的检索算法在不同规模数据库中,按照关键字检索方式检索相关数据所需时间,以及返回数据量大小进行对比。文献[10]中的检索算法利用双线性对对每一数据进行检索对比,检索计算量随着数据库规模的增加线性增加,而且该模型与其他现有模型相同仅简单的将匹配的所有密文数据返回至用户端,当数据库规模较大时,数据通信量较大对系统带宽要求提高,增加用户数据处理负担。新模型中 CS 按照 UDACM 中的逻辑规则仅就用户访问权限内的数据进行检索,随着数据库规模的增加检索计算量增加较小、检索效率高,而且新模型利用保序加密算法将加密后的关键字相关度作为索引属性,通过对匹配数据进行进行相关度安全比较,将相关度较高的前 k 项返回至用户,大大降低系统通信负担,提高检索结果的准确性,降低用户数据处理负担。

对于动态用户密钥管理,本文通过不同数目用户动态变化,统计分析了文献[12、15]设计的不同权限多用户密钥管理模型和本文提出的新模型中更新密钥的数目,见图 2。

图 2 反映了模型中每次用户动态变化密钥更新的数目,由图 2 可知在文献[12、15]中为保证数据安全性,当用户发生动态变化时需要更新与该用户相关的所有用户对应的密钥,因此当用户动态变化次数增加时,模型密钥管理难度大、效率低;本文设计的新模型中,对于用户的动态变化仅需对应更新

UACL 中相应代理加密密钥即可,与其他用户访问密钥无关,因此模型密钥管理灵活高效。

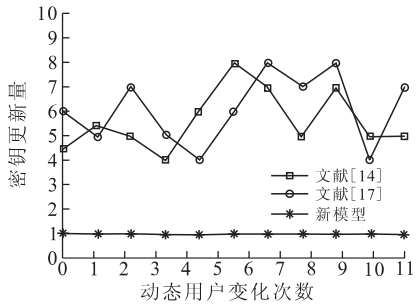

图 2　用户动态变化密钥更新量

Fig.2　Key update quantity for dynamic change of user

5　结语

本文在分析总结现有云存储模型关键技术的基础上,通过研究发现目前没有一个以云服务器为中心节点,能够支持密文检索的不同权限的多用户安全访问控制模型。针对该问题,利用椭圆曲线公钥密码算法设计具有语义安全的关键字检索算法,并将保序加密[9]后的关键字相关度作为密文索引的一部分,提高检索结果的准确性,同时降低系统带宽要求和用户数据处理负担。利用具有半诚实特点的云服务器按照用户访问控制列表(UACL)和用户——数据访问控制矩阵(UDACM)对用户检索请求进行二次代理加密,实现了不同权限的多用户访问控制,有效解决了用户动态变化所带来的密钥管理问题。通过对模型数据安全性和工作效率的对比分析,证明了新模型的可行性。

参考文献(References):

[1] Dawn Xiaodong Song, David Wagner, Adrian Perrig. Practical techniques for searches on encrypted data[C]//In the 2000 IEEE symposium on security and privacy,2000:44-55.

[2] Masayuki Yoshino, Ken Naganuma, Hisayoshi Satoh. Symmetric searchable encryption for database applications[C]//2011 international conference on network-based information systems. Tirana, Albania: IEEE computer society, 2011: 657-662.

[3] Boneh C, Crescenzo G D, Ostrovsky R, et al. Public key encryption with keyword search[C]//Pro int'l conf advances in cryptology (EUROCRYP'04). Berlin: Springer-verlag, 2004:506-522.

[4] Golle P, Staddon J, Waters B R. Secure conjunctive keyword search over encrypted data[C]//Proc Second int' conf applied cryptography and network security. Berlin: Springer-verlag, 2004:31-45.

[5] Hwang Y H, Lee P J. Public key encryption with conjunctive keyword search and its extension to a multi-user system[C]//Pairing-based cryptography. Berlin: Springer-verlag, 2007:31-45.

[6] Cong Wang, Ning Cao, Kui Ren, et al. Enabling secure and efficient ranked keyword search over outsourced cloud data[J]. IEEE transcantions parallel and distributed systems, 2012, 23(8): 1467-1479.

[7] Di Vimercati S D C, Foresti S, Jajodia S, et al. A data outsourcing architecture combining cryptography and access control[C]//Proc of ACM workshop on computer security architecture. Fairfax, VA: ACM SIGAC, 2007: 63-69.

[8] Yu S, Wang C, Ren K, et al. Achieving secure, scalable, and fine-grained data access control in cloud computing[C]//IEEE INFOCOM 2010. San Diego, CA: Vnited states institute of electrical and electronics engineers inc, 2010: 63-71.

[9] Wang W, Owens Z Li R, Bhargava B. Secure and efficient access to outsourced data[C]//Proc of ACM cloud computing security workshop. New York: ACM press, 2009: 55-65.

[10] Yang Yanjiang. Towards multi-user private keyword search for cloud computing[C]//2011 IEEE 4th international conference on cloud computing. Washington, DC: IEEE press, 2011: 758-759.

[11] Selim G Akl, Peter D Taylor. Cryptographic solution to a problem of access control in a hierarchy[J]. ACM transactions on computer systems, 1983, 1(3): 239-248.

[12] Morteza Nikooghadam, Ali Zakerolhosseini, Mohsen Ebrahimi Moghaddam. Efficient utilization of elliptic curve cryptosystem for hierarchical access control[J]. The journal of systems and software, 2010, 83: 1917-1929.

[13] Das M L, Saxena A, Gulati V P, et al. Hierarchical key management scheme using polynomial interpolation[J]. SIGOPS oper syst rev, 2005, 39(1): 40-47.

[14] Chung Yu Fang, Lee Hsiu Hui, Lai Feipei, et al. Access control in user hierarchy based on elliptic curve cryptosystem[J]. Information sciences, 2008, 178: 230-243.

[15] Jeng F G, Wang C M. An efficient key-management scheme for hierarchical access control based on elliptic curve cryptosystem[J]. Journal of systems and software, 2006, 79(8): 1161-1167.

[16] Koblitz N. Elliptic curve cryptosystems[J]. Mathematics of computation, 1987, 48(9): 203-209.

[17] Boneh D, Franklin M. Identity-based encryption from the weil pairing[C]//Advances in Cryptology-CRYPTO 2001, LNCS 2139, Berlin: Springer-verlag, 2001: 213-229.

[18] Dong Changyu, Russello Giovanni, Dulay Naranker. Shared and searchable encrypted data for untrusted servers [J]. Journal of computer security, 2008, 19(3): 127-143.

[19] Agrawal R, Kiernan J, Srikant R, et al. Order-preserving encryption for numeric data[C]//SIGMOD 2004, New York: ACM press, 2004: 563-574.

(编辑:徐楠楠)