doi:10.3969/j.issn.1000-436x.2013.z1.012

具有细粒度访问控制的隐藏关键词可搜索加密方案

杨旸[1,2]，林柏钢[1,2]，马懋德[1,3]

(1. 福州大学 数学与计算机工程学院，福建 福州 350108；
2. 福州大学 网络系统信息安全福建省高校重点实验室，福建 福州 350108；
3. 南洋理工大学 电子与电气工程学院，新加坡 新加坡 639798)

摘 要：针对现有的可搜索加密算法在多用户环境中密钥管理难度大并且缺乏细粒度访问控制机制的问题，利用基于密文策略的属性加密机制(CP-ABE, ciphertext-policy attribute based encryption)实现了对隐藏关键词可搜索加密方案的细粒度访问控制。数据拥有者可以为其在第三方服务器中存储的加密指定灵活的访问策略，只有自身属性满足该访问策略的用户才有权限对数据进行检索和解密。同时还能够实现对用户的增加与撤销。安全性分析表明方案不仅可以有效地防止隐私数据的泄露，还可以隐藏关键词的信息，使得第三方服务器在提供检索功能的同时无法窃取用户的任何敏感信息。方案的效率分析表明，该系统的检索效率仅为数十微秒，适合在大型应用系统中使用。

关键词：隐藏关键词检索；可搜索加密；细粒度访问控制；用户增加与撤销

中图分类号：TN918.1 **文献标识码**：A **文章编号**：1000-436X(2013)Z1-0092-09

Secure hidden keyword searchable encryption schemewith fine-grained and flexible access control

YANG Yang[1,2], LIN Bo-gang[1,2], MA Mao-de[1,3]

(1. College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China;
2. Key Lab of Information Security of Networks Systems of Fujian Province, Fuzhou University, Fuzhou 350108, China;
3. School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore)

Abstract: Existing searchable encryption schemes have difficulties in key management for multiple users and could not provide fine-grained access control mechanism. Aiming at solving these problems, a hidden keyword searchable encryption scheme with fine-grained access control was proposed utilizing CP-ABE (ciphertext-policy attribute based encryption) algorithm. Data owners allocate specific and flexible access policy on their data that is stored on a third-party data server. Only those users that has attributes satisfing the access policy are authorized to search encrypted data and decrypt returned results. Moreover, the suggested system has the function to add and revoke user. Security analysis shows that the scheme could not only prevent the leakage of private data but also hide the information of keywords. It deters a third-party storage provider from intercepting users' sensitive information when a search function is provided. The efficiency analysis shows that the efficiency of retrieval keeps no more than tens of microsecond and this scheme is suitable for large scale system.

Key words: hidden keyword search; searchable encryption; fine-grained access control; add and revoke user

收稿日期：2013-06-10

基金项目：国家自然科学基金资助项目(60970119, 61100231, 61103175, 61173151)；国家重点基础研究发展计划（“973”计划）基金资助项目（2007CB311201）

Foundation Items: The National Natural Science Foundation of China (60970119, 61100231, 61103175, 61173151); The National Basic Research Program of China (973 Program) (2007CB311201)

1　引言

1.1　背景介绍

随着电子医疗系统和网络云存储服务的高速发展，为了节省管理成本和快捷，很多企业用户及个人将大量数据共享至第三方数据中心存储。新兴服务的出现给用户带来便利的同时也伴随着严峻的安全隐患。这些数据中往往含有大量的企业或个人隐私信息，一旦泄露将会带来无法挽回的损失，如何保证敏感数据的隐私和机密性成为急需解决的棘手问题。加密技术是一种用密码学的方法来提供机密性和隐私的有效手段。将数据先做加密算法处理后，再上传至第三方服务器，可以防止数据被非法读取。然而，加密手段在提供保密性的同时，也给用户从海量数据中查询所需要的数据带来了困难。传统的加密技术将原始消息变为形同乱码的密文，在不进行解密的情况下，很难读出里面的明文信息，这与搜索所需的语义性是相矛盾的。因此，如何在非可信的环境下实现对加密数据的高效检索成为了最近的一个研究热点。

在传统的信息检索中，用户所关心的内容往往是希望对检索服务器以外的所有实体都是保密的，而不对服务器保密。因此，在用户查询数据的过程中，检索服务器可以无阻碍地获取用户所关心的任何内容。但事实上，服务器未必是可信的，数据查询和存储提供商可能会利用用户的敏感信息来做一些损害用户利益的事情。例如，电子医疗系统的存储服务器可能会将用户敏感的医疗信息泄露给保险公司，使得患病的病人无法像正常人一样为自己的健康投保。再如，金融投资者在使用云存储服务器查询公司的信息时，用户不希望任何人知道它在对哪些账目进行查询，如果信息泄露，将会对投资带来极为不利的影响。因此，用户不仅要防止敏感数据被他人获取，还要防止服务器窃取自己的隐私，这样才能实现用户的隐私安全，同时达到安全检索的目的。

可搜索加密技术(SE, searchable encryption)是一种新型的加密手段，在为文件或数据提供加密功能同时，还能够实现对加密后的密文进行检索的功能。在可搜索加密系统中，提供检索功能的第三方服务器无法通过监控用户的查询过程和用户存储的密文信息来获取与明文相关的任何信息。

1.2　相关工作

1992 年，Ostrovsky[1]最先提出设计一个系统，让用户自己加密数据并上传至远程服务器，同时对远程服务器保持数据的隐私性。2000 年，Song 等[2]提出对称可检索加密方案(SSE, symmetric searchable encryption)，该方案的检索操作仅需一次交互。随后又提出了一些新 SSE 方案[3,4]。2004 年 Boneh 等[5]开创性地提出带关键词检索的公钥加密方案(PEKS, public key encryption with keyword search)。该方案基于公钥密码学，用户在加密文件之前抽取出其中的关键词，用公钥对关键词进行加密，建立关键词索引(index)，然后将加密文件和关键词索引一同上传至服务器中。在文件检索过程中，为了保护被检索信息的隐私性，用户利用自己的私钥对所要查询的关键词生成陷门信息(trapdoor)。服务器接收到查询请求和陷门信息后，用匹配算法(test algorithm)实现对陷门信息和索引之间的配对查询，找出符合的条目(item)，从而返回相应的密文。随后一些学者提出了改进方案[6,7]。

2008 年，Baek 等[8]提出了不需要安全信道(secure channel free)的 PEKS 方案，解决了在向服务器传送关键词限门时需要建立安全信道的问题。Byun 等[9]和 Yau 等[10]用离线关键词攻击法对现有 PEKS 方案进行密码分析，并指出 8 个现有的可搜索加密方案都无法抵抗离线关键词攻击。2008 年，郑州信息工程大学的谷春香等[11]对 Beak 的无安全信道 PEKS 方案进行了改进，降低了服务器和用户的计算量，然而方案的安全性是基于随机预言机模型。Rhee 等[12]尝试找出一种通用的指定测试者(designated tester)的 PEKS 构造方案。

为了实现同时对多个关键词进行检索，Golle 等[13]于 2004 年首先提出连结关键词可搜索加密的概念(PECKS, publick-key encryption with conjunctive keyword search)并建立了安全模型。在连结关键词可搜索加密方案中，每个加密文件都与若干个关键词相关联，用户可以对多个关键词生成一个陷门来进行检索，服务器只需通过一次测试操作即可搜索出包含多个关键词的加密文件。与单关键词可搜索加密方案相比，连结关键词可搜索加密方案效率更高。Park 等[14]构造了 2 个支持多关键词(multi-keyword)检索的公钥可搜索加密方案。然而，这些系统都仅仅适用于用户将数据加密后存储至第三方服务器，然后用户自身对加密数据进行查询的应

用场景。无法支持另一个用户对数据进行秘密检索，除非用户将私钥完全泄露给第三方，这显然并不适用与大型网络环境。随后，文献[15~17]中又提出了改进方案以提高算法效率和对用户的隐私性保护。

2006 年，Curtmola 等[18]通过广播加密技术将可搜索加密方案扩展至多用户情形。授权用户的集合共享密钥，只有拥有密钥的用户才能检索数据。然而，该方案中的加密数据库不便于更新。2007 年，Hwang 等[19]构造了一个连结关键词公钥可搜索加密方案并将其扩展成多用户方案，2 个方案都是在随机预言机模型下证明是安全的。2008 年，Bao 等[20]构造了多用户可搜索加密方案，并给出随机预言机模型下的安全性证明；方案中多个用户共享对称密钥，无法实现多用户安全地动态加入和退出。随后又提出一些新的能支持多用户搜索的加密系统[21~23]，但是仍然无法对用户的访问权限进行细粒度的管理。

访问控制机制可以用来对资源进行保护，防止未经授权的实体访问特定的信息或已授权用户非法访问超权限数据。根据不同的访问控制策略，访问控制机制主要可分为：自主访问控制(DAC, discretionary access control)、强制访问控制(MAC, mandatory access control)、基于角色访问控制(RBAC, role-based access control)和基于属性的访问控制(ABAC, attribute-based access control)。在自主访问控制机制(DAC)中，数据所有者或具有管理权限的主体利用访问控制列表枚举出所有具有访问权限的客体。在大型系统中，客体的数量将异常巨大。在某些应用环境(如云存储)中，数据拥有者难以在上传数据时就确定能够访问各种资源的所有人员名单。强制访问控制(MAC)系统为不同用户分配安全等级，并且规定信息只能由低等级向高等级流动，过多的限制导致了使用的不灵活性，授权的可管理性也较差。基于角色的访问控制(RBAC)将用户与角色关联，然后将角色与权限挂钩，实现了用户、角色和访问权限之间的映射关联，对访问权限的管理更加灵活。基于属性的访问控制(ABAC)是对 RBAC 的拓展，将不同属性(或属性集合)赋予不同的主体，便于处理分布式环境下的细粒度授权访问和大规模用户的动态管理。

文献[24]利用基于密文策略的属性基加密(CP-ABE)设计了一种代理重加密方法(HCRE, hybrid cloud re-encryption)来实现对密文的访问控制。但是该方案并没有提供对密文进行关键词检索这一重要功能。此外，对用户的权限进行撤销时，需要更改所有文件的访问控制树(access tree)，这样的更新在大规模的系统中将会带来较大的计算开销。文献[25]利用代理重加密方法来对多用户的访问权限进行管理，但无法对文件的访问控制权限的细粒度进行控制。并且针对不同用户的检索请求，服务提供商都需要将数据和目录用不同用户的辅助密钥重新加密一次，再供该用户搜索，这将大大增加服务器的运算量和服务提供时延。

1.3 设计中的挑战

现有的可搜索加密方案中[18~23]对多用户系统的管理还有很多的缺陷。通常采用多用户共享密钥的方式来实现数据和检索权限的共享。然而数据的加密密钥被多个用户共享会大大增加其泄露的可能性。若要取消某个用户的访问权限，只能更换密钥，这样所有用该密钥加密的数据都必须被加密，然后用更新后的密钥加密，再上传至服务器，这种方式显然会带来大量的传输开销、计算开销以及时间开销。

此外，无法对加密数据的访问权限进行细粒度的管理。在很多应用场景中，数据拥有者并不希望把对所有数据解密的权限交付给任意一个用户，而是希望部分具有特定需求的用户访问某些其所需要的那一部分数据，其他与用户业务上无直接关系的数据则是保持秘密性的。例如，在电子医疗系统中，病人可能希望为其进行医治的骨科大夫只能访问用户与骨科相关的医疗信息，而病人带有传染病性质的数据对于骨科大夫而言是不可见的。这就需要对用户的访问权限进行更为细致的划分。此外，系统还应该便捷地实现对用户的管理，例如用户的增加和访问权限的召回(revocation)。

1.4 本文的贡献

针对现有的多用户系统存在的问题，本文设计了具有细粒度访问控制机制的关键词检索系统，其优势列举如下。

1) 密钥管理灵活。在本文所设计系统中，每个用户拥有自身的一组身份属性，这些属性与用户的私钥相关联。即不同的用户拥有独立的私钥，可以避免多用户共享密钥而带来的泄露风险。

2) 细粒度的访问控制。数据拥有者(data owner)在上传文件之前，为不同的文件设置不同的访问策略(policy)。在加密过程中由访问策略来进行控制，

使得访问策略嵌入到密文中，只有当用户属性满足密文的访问策略时，才拥有检索权限并能解密相应的密文。

3) 数据保密性。数据拥有者对数据加密上传至服务器后，没有访问权限的用户和服务器均无法读取密文中的私有信息。

4) 关键词隐私性。在服务器为用户提供检索操作的过程中，用户所检索的关键词信息是以陷门的形式提交，文件中所存储的关键词索引信息也是以密文的方式存放，服务器无法在此过程中窃听到任何与关键词有关的明文信息。

5) 抗共谋性。如果一组用户中没有任何一个人有权限访问某些隐私数据，那么它们即使将自己的密钥组合在一起也无法实现解密该数据的目的。

2　系统模型

在如图 1 所示的多用户可搜索系统中，含有以下几个实体。

1) 密钥分发中心。密钥分发中心在本系统中是一个可信的第三方，负责为系统生成公开参数，并为用户生成和分发密钥。根据数据拥有者的请求增加或撤销用户的访问权限，并向云存储中心提供用户撤销证书。

2) 云存储中心。云存储中心是诚实但是好奇(honest but curious)的第三方数据中心。它主要负责数据的存储和检索，会诚实地根据用户的检索请求返回相应的加密数据，但是也会在通信过程中试图窃取用户的隐私。

3) 数据拥有者。数据拥有者在上传文件前，为不同的文件设置不同的访问策略(policy)。对要上传至云服务器的数据进行加密处理，并且在加密过程中由访问策略来进行控制，使得访问策略嵌入到密文中。拥有增加或召回其文件访问者的管理权限。

4) 数据访问者。数据访问者拥有一些指定的属性集合以及相应的密钥。可以向云存储中心提供查询请求，若其属性满足加密文件的访问树，则可以访问该数据。

3　预备知识

3.1　双线性映射

定义 1　双线性映射。令 G 为阶数为素数 p 的加法循环群，G_1 为具有相同阶次的乘法循环群。假设 G 和 G_1 中的离散对数问题是困难问题。双线性映射是指满足下列性质的一个映射。

$$\hat{e}: G\times G \to G_1 。$$

1) 双线性性。对任意 $P,Q\in G$, $a,b\in \mathbb{Z}_p^*$, 满足 $\hat{e}(P^a,Q^b)=\hat{e}(P,Q)^{ab}$ 。

2) 非退化性。存在 $P,Q\in G$, 使得 $\hat{e}(P,Q)\neq 1$ 。

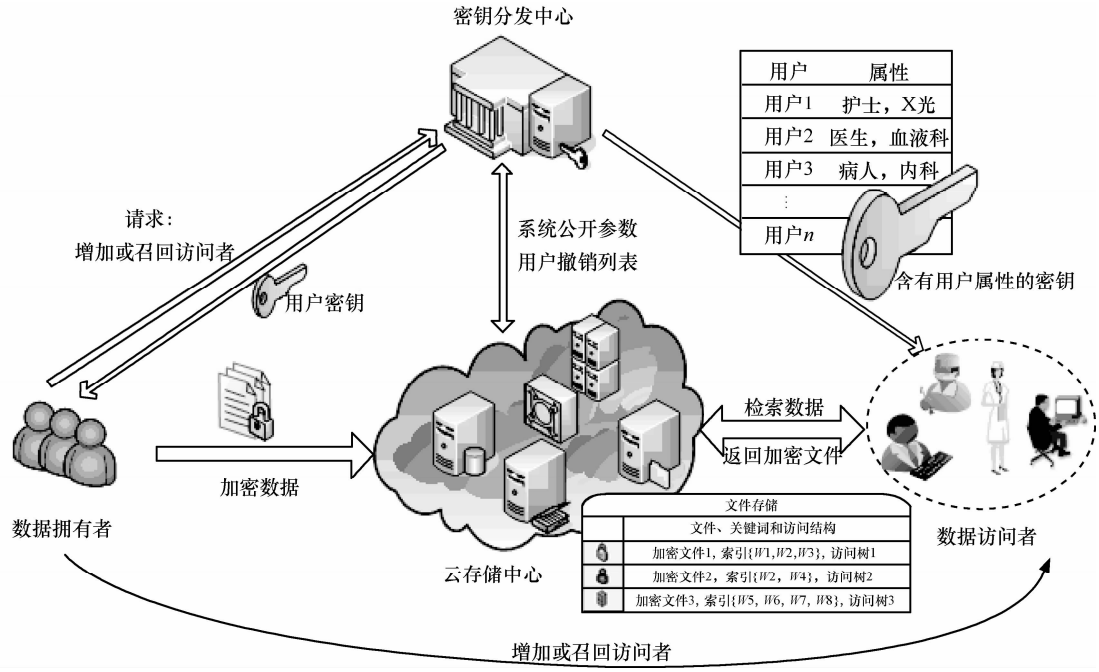

图 1　系统模型

3) 可计算性。对所有的 $P,Q\in G$，存在有效的算法计算 $\hat{e}(P,Q)$。

3.2 访问结构与访问树

访问结构(access structure)是一个定义了授权访问子集和非授权访问子集的概念，用来描述访问控制的策略。其工作原理类似于秘密共享(secret sharing)，访问结构定义了一些参与者的集合可以重构秘密以及另一些参与者的集合不可以重构秘密。

设参与者集合为 $P=\{p_1,p_2,\cdots,p_l\},l\in Z$，共享的秘密为 s。能重构秘密 s 的参与者子集称为授权子集；不能重构秘密 s 的参与者子集称为非授权子集。所有授权子集构成的集合称为访问结构，用符号 Γ 表示。所有非授权子集构成的集合表示为 $\bar{\Gamma}=2^P\setminus\Gamma$（$2^P\setminus\Gamma$ 表示所有在 2^P 中而不在 Γ 中元素组成的集合，2^P 表示属性集合 P 的所有子集)。

通常用访问树来表示一个访问结构的逻辑表示，即访问策略。令树 T 表示一棵访问树，每一个非叶子节点代表一个门限门(threshold gate)，由它的孩子节点(children)和门限值(threshold value)表示。用 num_x 表示节点 x 的孩子节点的个数，k_x 表示节点 x 的门限值，则 $1\leqslant k_x\leqslant num_x$。当 $k_x=num_x$ 时，节点 x 的逻辑门是与门(AND)，表示当所有属性都满足时，才能恢复秘密；当 $k_x=1$ 时，节点的逻辑门是或门(OR)，表示所有属性中只要满足一个就可以恢复秘密。访问树 T 的每个叶子节点 l 都与一个属性 $attr_l$ 相关联，$parant(l)$ 表示叶子节点 l 在访问树 T 中的父节点。每个父节点 x 的孩子节点被标记为 1~num_x 之间的某个数，用 $index(x)$ 表示这个标号(相对于节点 x)。

令 T_x 表示 T 的子树，该子树的根节点为 x。若一个属性集合 Ψ 满足访问树 T_x 的访问条件，则表示为 $T_x(\Psi)=1$。若不满足，则 $T_x(\Psi)=0$。判断属性集合 Ψ 是否满足访问树 T_x 的访问条件(即计算 $T_x(\Psi)$)需要用递归的方式计算。

1) 若 x 不是叶子节点，为节点 x 的每一个孩子节点 x' 计算 $T_{x'}(\Psi)$。当且仅当至少有 k_x 个 $T_{x'}(\Psi)$ 等于 1 时，返回 1；否则返回 0。

2) 若 x 是叶子节点，当且仅当 $attr_x\in\Psi$ 时，返回 1；否则返回 0。

例如图 2 所示一个访问控制树的结构。爱丽丝表示电子医疗记录系统的使用者和数据拥有者。根据它所设置的文件访问树能访问该文件的有 3 类人：第 1 类是爱丽丝本人；第 2 类是同时含有“骨科”和“主治医生”2 种属性的用户；第 3 类是同时拥有“2013 年”、“X 光”和“护士长”3 种属性的用户。若用户拥有的属性为“X 光”和“实习医生”，则无法访问该数据。

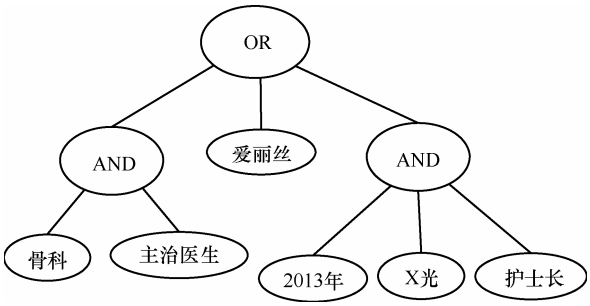

图 2 访问控制树

3.3 基于密文策略的属性加密

基于属性的加密系统把访问结构的概念融入到基于身份的加密系统中，在用户的私钥或密文中引入一个访问结构，使得具备某些属性(授权访问属性)的用户可以解密密文，而属性不满足的用户则无法解密密文。

本文是利用基于密文策略的属性机密技术(CP-ABE, ciphertext-policy attribute based encryption)[26]来实现细粒度的访问控制系统。CP-ABE 算法主要包含以下 4 个算法[27~29]。

$Setup(k)$：初始化算法的输入为安全参数 k，算法生成系统公开参数(global parameter) GP 和主密钥(master key) K_{MK}。

$KeyGen(K_{MK},\Psi)$：密钥生成算法的输入为公开参数 GP 和属性集合 Ψ，算法输出与属性集合 Ψ 相关联的密钥 SK。

$Encrypt(GP,M,\mathbb{A})$：加密算法的输入为系统公开参数 GP、消息 M 和访问结构 $\mathbb{A}$，算法输出消息 M 的密文 CT 使得只有属性集合满足访问结构 $\mathbb{A}$ 的用户才能解密 CT。

$Decrypt(GP,CT,SK)$：解密算法的输入为系统公开参数 GP、密文 CT (与访问结构 $\mathbb{A}$ 关联)和用户密钥 SK (与属性集合 Ψ 关联)。若属性集合 Ψ 满足访问结构 $\mathbb{A}$，则算法解密密文 CT 并返回消息 M。

3.4 具有细粒度访问控制的可搜索加密算法模型

首先定义算法模型中涉及的符号。令 $U=\{U_1,U_2,\cdots,U_l\}$ 表示用户的集合，$\Phi=\{\varphi_1,\varphi_2,\cdots,\varphi_n\}$ 表示属性的集合，Ψ_i 表示用户 U_i 的属性集合，$\mathbb{A}$ 表示密文

的访问控制策略。利用基于密文策略的属性加密来构造具有细粒度访问控制的可搜索加密算法由以下几个算法组成。

GlobalSetup(k)：密钥分发中心(key distribution center)输入安全参数 k，算法输出系统公开参数(global parameter) GP 和主密钥(master key) K_{MK}。

KeyGen(GP, K_{MK}, U_i, Ψ_i)：密钥分发中心输入公开参数 GP、主密钥 K_{MK}、用户的唯一身份标识 U_i 和用户的属性集合 Ψ_i，算法输出用户的公钥 pk_{U_i}，私钥 sk_{U_i} 和辅助密钥(complementary key) ck_{U_i}。密钥分发中心将 (ck_{U_i}, sk_{U_i}) 发送给用户，用户将 sk_{U_i} 私藏。密钥分发中心将 (U_i, ck_{U_i}) 发送给数据中心存储。

Encrypt($M, W, pk_{U_i}, \mathcal{A}$)：数据拥有者(data owner) U_i 输入明文 M 及关键词集合 $W=\{w_1, w_2, \cdots\}$、公钥 pk_{U_i} 和访问结构 $\mathcal{A}$，算法输出明文 M 的密文 CT 和关键词集合 W 的私密索引 $I(W)$，数据拥有者将 $\{CT, I(W), \mathcal{A}\}$ 上传至第三方存储服务器。

Trapdoor(sk_{U_i}, W)：用户检索含有关键词集合 W 的所有密文，需要生成关键词 W 的陷门。算法输入用户 U_i 的私钥 sk_{U_i} 和关键词 W，输出陷门 T_W。

Retrieve($U_i, CT, I(W), T_W, \Psi_i, \mathcal{A}$)：第三方存储服务器收到用户的检索请求后，搜索所有与关键词陷门 T_W 匹配的索引 $I(W)$，然后验证用户 U_i 的属性集合 Ψ_i 是否满足与 $I(W)$ 相对应密文 C 的访问结构 $\mathcal{A}$，若满足则返回索引匹配且满足访问权限的密文集合 $CT=\{CT_1, CT_2, \cdots\}$。

Decrypt($CT, sk_{U_i}, \Psi_i, \mathcal{A}$)：用户收到返回的密文集合 CT 后，用自己的私钥 sk_{U_i} 解密出明文集合 $M=\{M_1, M_2, \cdots\}$。

Revoke(K_{MK}, U_i)：若用户 U_i 的访问权限到期或提前撤回其访问权限，则密钥分发中心生成撤销证书 $Revoke_{U_i}=\{U_i, Date, Sig_{K_{MK}}(U_i, Date)\}$，证书中包括用户名、撤销时间以及密钥分发中心用主密钥对撤销信息的签名。密钥分发中心将撤销证书发送给第三方存储服务器。存储服务器删除存储的 (U_i, ck_{U_i}) 条目，并将该证书放入撤销链表中。被撤销权限的用户进行访问时，若该用户在撤销链表中，则存储服务器拒绝其检索请求。

4　具有细粒度访问控制的可搜索加密方案

本文提出的可搜索加密算法基于 CP-ABE 技术，但是功能方面又有很多不同点。传统的 CP-ABE 算法只包含 4 个算法，仅当用户的属性满足特定的访问结构时，才能解密相应的明文。但是 CP-ABE 算法不能为用户提供对密文信息进行检索的功能。当海量数据存储在远程服务器上时，CP-ABE 体系结构中的用户如果要检索包含一些指定关键词的文件，则需要对其拥有访问权限的所有文件进行解密，然后对明文数据进行检索。显然，这将耗费大量的通信开销和计算开销。然而，具有细粒度访问控制的可搜索加密体制含有 7 个算法，可以让用户在不对密文信息进行解密的前提下检索含有特定关键词(或关键词集合)的信息，同时对用户检索和解密的权限进行细粒度的管理，此外还能对用户的访问权限进行撤销。

利用CP-ABE方案[30]构造出具有细粒度访问控制策略的可搜索加密方案。令 $\Phi=\{\varphi_1, \varphi_2, \cdots, \varphi_n\}$ 表示用户的属性集合，散列函数 $H:\{0,1\}^* \to G$，方案的具体步骤如下所示。

GlobalSetup(k)：密钥分发中心输入安全参数 k，算法选择阶为 p 的群 G，生成元 g 和随机元素，$\alpha, t_1, t_2, \cdots, t_n \in Z_p^*$。计算 $y=\hat{e}(g,g)^{\alpha}$ 以及 $T_i=g^{t_i}$。公开参数为 $GP=(g, y, T_i)$，主密钥为 $K_{MK}=(\alpha, t_i)$，其中，$1 \leqslant i \leqslant n$。

KeyGen(GP, K_{MK}, U_i, Ψ_i)：密钥分发中心输入公开参数 GP、主密钥 K_{MK}、用户的唯一身份标识 U_i 和用户的属性集合 Ψ_i，算法选择随机数 $r, \beta_i, \eta \in Z_p^*$，并计算 $d_0=g^{\alpha-r}$, $d_0'=\eta$, $ck_{U_i}=\beta_i d_0'$。对于 Ψ_i 中的每一个属性 φ_j，计算 $d_j=g^{rt_j^{-1}}$。用户 U_i 的私钥为 $sk_{U_i}=(d_0, d_0', d_j \mid \forall \varphi_j \in \Psi_i)$。密钥分发中心将 (ck_{U_i}, sk_{U_i}) 发送给用户，用户将 sk_{U_i} 私藏。密钥分发中心将 (U_i, ck_{U_i}) 发送给数据中心存储。

Encrypt($M, W, pk_{U_i}, \mathcal{A}$)：数据拥有者(data owner) U_i 输入明文 M 及关键词集合 $W=\{w_1, w_2, \cdots\}$、公钥 pk_{U_i} 和访问结构 $\mathcal{A}$，算法计算如下。

1) 选择随机数 $s \in Z_p^*$ 并计算 $C_0=g^s, C_1=My^s$。

2) 令访问结构 $\mathcal{A}$ 对应的访问树为 T，根节点为 R，根节点 R 的值为 s。根据访问树 T 的节点间“与”“或”关系做如下分配：①当父节点与子节点之间的关系为“或”结构时，其子节点的值也为 s。②当父节点与子节点之间的关系为“与”结构时，设其子节点的个数为 t，随机选择

$s_1, s_2, \cdots, s_{t-1} \in Z_p^*$，计算 $s_t = s - \sum_{i=1}^{t-1} s_i$，将 $s_1, s_2, \cdots, s_t$ 分别赋值给这 t 个子节点。

3) 对于 T 的每个叶子节点 $\varphi_{j,i} \in T$，计算 $C_{j,i} = T_j^{s_i}$。明文 M (访问结构为 $\mathcal{A}$，访问树为 T)的密文 $CT = (T, C_0, C_1, C_{j,i} \mid \varphi_{j,i} \in T)$。

4) 随机选取元素 $\tau \in Z_p^*$，计算 $A = g^{\tau}$，$B = \hat{e}(H(W), g^{\tau \cdot ck_{U_i}})^{1/d_0'}$，关键词集合 W 的私密索引为 $I(W) = (A, B)$。

算法输出明文 M 的密文 CT 和关键词集合 W 的私密索引 $I(W)$，数据拥有者将 $\{CT, I(W), \mathcal{A}\}$ 上传至第三方存储服务器。

$Trapdoor(sk_{U_i}, W)$：用户检索含有关键词集合 W 的所有密文，需要生成关键词 W 的陷门。算法输入用户 U_i 的私钥 sk_{U_i} 和关键词 W，随机选取元素 $\lambda \in Z_p^*$，输出陷门 $T_W = (T_{W,1}, T_{W,2}) = (\lambda, H(W)^{\lambda/d_0'})$。

$Retrieve(U_i, CT, I(W), T_W, \Psi_i, \mathcal{A})$：第三方存储服务器收到用户的检索请求后，搜索所有与关键词陷门 T_W 匹配的索引 $I(W)$，即验证等式 $\hat{e}(T_{W,2}, A^{ck_{U_i}}) = B^{T_{W,1}}$ 是否成立。若成立则为匹配项。等式成立的原因为

$$\hat{e}(T_{W,2}, A^{ck_{U_i}}) = \hat{e}(H(W)^{\lambda/d_0'}, (g^{\tau})^{\beta_i d_0'}) = \hat{e}(H(W), g)^{\lambda \tau \beta_i}$$

$$B = \hat{e}(H(W), g^{\tau \cdot ck_{U_i}})^{\lambda/d_0'} == \hat{e}(H(W), g)^{\lambda \tau \beta_i}$$

然后验证用户 U_i 的属性集合 Ψ_i 是否满足与 $I(W)$ 相对应的密文 C 的访问结构 $\mathcal{A}$，若满足则返回索引匹配，且满足访问权限的密文集合为 CT={CT_1, CT_2, $\cdots$}。

$Decrypt(CT, sk_{U_i}, \Psi_i, \mathcal{A})$：用户收到返回的密文集合 CT 后，用自己的私钥 sk_{U_i} 解密。首先，选择满足访问结构 $\mathcal{A}$ 的最小属性集合 $\Psi_i' \in \Psi_i$，通过如下计算即可恢复出明文 M。

$$\prod_{\varphi_j \in \Psi_i'} \hat{e}(C_{j,i}, d_j) = \prod_{\varphi_j \in \Psi_i'} \hat{e}(g^{t_j s_i}, g^{r t_j^{-1}}) = \hat{e}(g, g)^{rs}$$

$$\hat{e}(C_0, d_0) \cdot \hat{e}(g, g)^{rs} = \hat{e}(g^s, g^{\alpha - r}) \cdot \hat{e}(g, g)^{rs} = \hat{e}(g, g)^{\alpha s}$$

$$C_1 / \hat{e}(g, g)^{\alpha s} = M \cdot y^s / \hat{e}(g, g)^{\alpha s} = M$$

$Revoke(K_{MK}, U_i)$：若用户 U_i 的访问权限到期或提前撤回其访问权限，则密钥分发中心生成撤销证书 $Revoke_{U_i} = \{U_i, Date, Sig_{K_{MK}}(U_i, Date)\}$，证书中包括用户名、撤销时间以及密钥分发中心用主密钥对撤销信息的签名。密钥分发中心将撤销证书发送给第三方存储服务器。存储服务器删除存储的 (U_i, ck_{U_i}) 条目，并将该证书放入撤销链表中。被撤销权限的用户进行访问时，若该用户在撤销链表中，则存储服务器拒绝其检索请求。

5 安全性分析

1) 数据保密性

在本文所提出的关键词可搜索方案中，文件在上传至服务器以前，明文数据已经在访问控制策略的控制下被加密成密文。因此服务器和恶意攻击者都无法获取用户数据的明文信息。

文件中的关键词提取是由数据拥有者在本地完成的，通过用户的私钥 sk_{U_i} 和辅助密钥 ck_{U_i} 加密成关键词索引：$I(W) = (g^{\tau}, \hat{e}(H(W), g^{\tau \cdot ck_{U_i}})^{1/d_0'})$。尽管辅助密钥 ck_{U_i} 也同时被分配给数据服务器，但是由于 sk_{U_i} 是由用户私藏，存储服务器依然无法读出索引中的关键词信息。因此，只要用户妥善保存私钥 sk_{U_i}，就可以保证数据明文和关键词信息的保密性。

此外，随机数 $\tau \in Z_p^*$ 的引入还可以有效地抵抗恶意攻击者的重放攻击。

2) 查询请求隐私性

用户在进行数据检索查询时，首先生成所要查询的关键词信息，然后利用自己的私钥对关键词的散列值进行运算生成陷门 $T_W = (T_{W,1}, T_{W,2}) = (\lambda, H(W)^{\lambda/d_0'})$。第三方存储服务器在接收到检索请求后，从关键词的陷门 T_W 无法获取 W 的明文信息，保证了查询请求的隐私性。此外，随机元素 $\lambda \in Z_p^*$ 也是为了防止查询请求的重放攻击，以阻止攻击者通过搭线窃听的方式截获陷门信息，然后重新发送给存储服务器，进行非法的查询操作。

3) 查询请求不可否认性

本系统支持多用户的查询请求。在传统的多用户系统中，多个用户共享相同的查询密钥。对于一个查询请求，数据服务器和密钥生成中心都无法判别出该请求是由哪个用户产生的，因此不具有查询请求不可否认性。而在本文所设计的多用户系统中，每个的用户都被分发了不同的密钥，因此查询请求所提交的 T_W 中含有了用户的私钥信息。任何其

他用户或存储服务器都无法伪造出该用户所生成的陷门信息 T_W，该用户也不可否认自己所生成的 T_W 是出于伪造。因此查询请求具有不可否认性。

4) 抗共谋性

由于消息密文的生成过程采用了基于密文的属性加密(CP-ABE)机制，一组用户中若都不具有访问某些隐私数据的权限，那么即使它们将自己的密钥组合也依然无法进行访问。因为只有在得到访问树中所有叶子节点对应的子秘密 $(C_{j,i} \mid \varphi_{j,i} \in T)$ 之后，才能恢复出：$\prod_{\varphi_j \in \Psi_i'} \hat{e}(C_{j,i}, d_j) = \prod_{\varphi_j \in \Psi_i'} \hat{e}(g^{t_j s_i}, g^{r t_j^{-1}}) = \hat{e}(g,g)^{rs}$，从而解密出明文 M。

6　方案比较与测试

将本文所提出的方案与其他几个典型的方案[19~21, 23~25]做比较，如表 1 所示。文献[24]提出的访问控制方案实现了动态地访问权限管理，但是并没有提供密文检索功能，并且撤销用户的访问权限时需要重新加密所有的密文，产生了较大地计算开销。文献[25]中的方案为多用户提供了可搜索加密方案，主要基于代理重加密技术，为不同用户提供检索服务时，都需要将关键词索引重新加密，对于大规模系统极为不便，该方案也无法对用户的访问权限进行细粒度的控制。文献[19~21,23]均提供了密文检索功能，但都无法实现细粒度的访问控制。此外，文献[19, 23]不能撤销已授权用户的访问权限。

表 1　　与经典方案的比较

方案	密文可检索	细粒度的访问控制	索引重加密	访问权限撤销	重加密密文(权限撤销时)
文献[19]	√	×	否	×	—
文献[20]	√	×	否	√	否
文献[21]	√	×	否	√	否
文献[23]	√	×	否	×	—
文献[24]	×	√	否	√	是
文献[25]	√	×	是	√	否
本文	√	√	否	√	否

本文提出的方案既提供了密文检索功能，又能实现基于属性的细粒度访问控制。无需对索引和密文进行重加密就可供具有不同属性的用户检索和访问，但前提是用户的属性与密文文件的访问控制树相符。

本文的方案利用基于双线性对的密码函数库(PBC, pairing-based cryptography)[31]进行仿真测试。作者选用的是密码库中提供的 A 类(type-A)椭圆曲线，群的阶数为 160 bit，相当于 1 024 bit 的离散对数运算强度。实验平台为个人台式电脑，中央处理器为 Pentium Dual core CPU(3.3 GHz)，内存为 2.0 GB，运行的操作系统为 Windows XP。

选择的实验参数为：系统中用户人数 100 人；每个文件抽取出的关键词个数至多为 10 个；用户属性总个数为 25；每个用户拥有的属性至多为 10 个；每颗访问树的叶子节点至多为 10 个。

经测试，GlobalSetup 算法运行时间至多为 0.941 s；为每一个用户生成密钥的算法 KeyGen 运行时间至多为 0.101 s；加密算法 Encrypt 运算时间至多为 0.325 s；生成陷门的算法 Trapdoor 运行时间至多为 0.037 s；检索验证算法的 Retrieve 执行时间至多为 0.074 s；解密算法 Decrypt 的执行时间至多为 0.223 s。

由此可见，该系统各个算法在普通 PC 机上的执行时间均不超过 1 s。在分布式系统环境以及云存储中，影响可搜索加密效率的最关键因素为检索阶段所需的时间，即寻找与用户关键词匹配的文件所需的时间。通过性能测试可以看出，本文方案的该项指标仅为几十毫秒，适合应用在大规模的数据存储系统中。

7　结束语

可搜索加密技术是一种能够同时提供加密和对密文检索功能的新型技术。针对现有可搜索加密系统所存在的多用户系统检索权限管理困难问题，设计了具有细粒度访问控制权限的关键词检索系统，可以根据用户的属性和文件访问控制策略的设定，对不同保密文件的访问权限进行管理。对提出的算法进行了安全性分析，表明其可以保证数据和关键词的隐私性，并且具有抗抵赖性、抗共谋攻击以及抗重放攻击的优点。对本文算法的性能仿真测试表明，系统具有较高的执行效率。

参考文献：

[1] OSTROVSKY R. Software Protection and Simulation on Oblivious RAMS[D]. MIT Ph D Thesis, 1992.

[2] SONG D, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[A]. Proceedings of IEEE Symposium on Security and Privacy[C]. 2000.1081-6011.

[3] LU H, GU D, JIN C, *et al*. Reducing extra storage in searchable sym-

metric encryption scheme[A]. Proceedings of the 4th International Conference on Cloud Computing Technology and Science[C]. 2012. 255-262.

[4] KAMARA S, PAPAMANTHOU C. Parallel and dynamic searchable symmetric encryption[EB/OL]. http://131.107.65.14/en-us/um/people/senyk/pubs/psse.pdf, 2013.

[5] BONEH D, CRESCENZO G, OSTROVSKY R, *et al*. Public key encryption with keyword search[A]. Proceedings of Eurocrypt[C]. 2004. 506-522.

[6] ZHANG R, IMAI H. Combining public key encryption with keyword search and public key encryption[J]. IEICE Transactions on Information and Systems, 2009, 92(5):888-896.

[7] LONG B, GU D, DING N, *et al*. On improving the performance of public key encryption with keyword search[A]. Proceedings of International Conference on Cloud and Service Computing[C]. 2012. 143-147.

[8] BAEK J, SAFAVI-NAINI R, SUSILO W. Public key encryption with keyword search revisited[A]. Proceedings of International Conference on Computational Science and Its Applications[C]. 2008. 1249-1259.

[9] BYUN J, RHEE H, PARK H, *et al*. Off-line keyword guessing attacks on recent keyword search schemes over encrypted data[A]. Proceedings of SIAM Conference on Data Mining[C]. 2006. 75-83.

[10] YAU W, HENG S, GOI B. Off-line keyword guessing attacks on recent public key encryption with keyword search schemes[A]. Proceedings of Autonomic and Trusted Computing[C]. 2008. 100-105.

[11] GU C, ZHU Y, PAN H. Efficient public key encryption with keyword search schemes from pairings[A]. Proceedings of Information Security and Cryptography[C]. 2008. 372-383.

[12] RHEE H, PARK J, LEE D. Generic construction of designated tester public-key encryption with keyword search[J]. Information Sciences, 2012, 205:93-109.

[13] GOLLE P, STADDON J, WATERS B. Secure conjunctive search over encrypted data[A]. Proceedings of Applied Cryptography and Network Security[C]. 2004. 31-45.

[14] PARK D, KIM K, LEE P. Public key encryption with conjunctive field keyword search[A]. Proceedings of Information Security Applications [C]. 2004. 73-86.

[15] YANG C, ZHANG W, XU J, *et al*. A fast privacy-preserving multi-keyword search scheme on cloud data[A]. Proceedings of International Conference on Cloud and Service Computing[C]. 2012. 104-110.

[16] DING M, GAO F, JIN Z, *et al*. An efficient public key encryption with conjunctive keyword search scheme based on pairings[A]. Proceedings of IEEE International Conference on Network Infrastructure and Digital Content[C]. 2012. 526-530.

[17] CHEN Z, WU C, WANG D, *et al*. Conjunctive keywords searchable encryption with efficient pairing, constant ciphertext and short trapdoor[A]. Proceedings of Intelligence and Security Informatics[C]. 2012. 176-189.

[18] CURTMOLA R, GARAY J, KAMARA S, *et al*. Searchable symmetric encryption: improved definitions and efficient constructions[A]. Proceedings of ACM Conference on Computer and Communications Security[C]. 2006. 79-88.

[19] HWANG Y, LEE P. Public key encryption with conjunctive keyword search and its extension to a multi-user system[A]. Proceedings of Pairing-based Cryptography[C]. 2007. 2-22.

[20] BAO F, DENG R, DING X, *et al*. Private query on encrypted data in multi-user settings[A]. Proceedings of International Conference on Information Security Practice and Experience[C]. 2008. 71-85.

[21] YANG Y, LU H, WENG J. Multi-user private keyword search for cloud computing[A]. Proceedings of IEEE International Conference on Cloud Computing Technology and Science[C]. 2011.264-271.

[22] DONG C, RUSSELLO G, DULAY N. Shared and searchable encrypted data for untrusted servers[J]. Journal of Computer Security, 2010, 19(3):367-397.

[23] LIU Q, TAN C, WU J, *et al*. Cooperative private searching in clouds[J]. Journal of Parallel and Distributed Computing, 2012, 72(8):1019-1031.

[24] 洪澄, 张敏, 冯登国. 面向云存储的高效动态密文访问控制方法[J]. 通信学报, 2011, 32(7):125-132.
HONG C, ZHANG M, FENG D G. Achieving efficient dynamic cryptographic access control in cloud storage[J]. Jounal on Communications, 2011, 32(7):125-132.

[25] 王映康, 罗文俊. 云存储环境下多用户可搜索加密方案[J]. 电信科学, 2012, 28(11):103-107.
WANG Y K, LUO W J. A scheme of multi-user searchable encryption in cloud storage[J]. Telecommunications Science, 2012, 28(11):103-107.

[26] BETHENCOURT J, SAHAI A, WATERS B. Ciphertext-policy attribute-based encryption[A]. Proceedings of IEEE Symposium on Security and Privacy[C]. 2007. 321-334.

[27] WATERS B. Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization[A]. Proceedings of Public Key Cryptography[C]. 2011. 53-70.

[28] HOHENBERGER S, WATERS B. Atribute-based encryption with fast decryption[A]. Proceedings of Public Key Cryptography[C]. 2013. 162-179.

[29] LEWKO A, WATERS B. New proof methods for attribute-based encryption: achieving full security through selective techniques[A]. Proceedings of Crypto[C]. 2012. 180-198.

[30] IBRAIMI L, TANG Q, HARTEL P, *et al*. Efficient and provable secure ciphertext-policy attribute-based encryption schemes[A]. Proceedings of Information Security Practice and Experience[C]. 2009. 1-12.

[31] LYNN B. The PBC library[EB/OL]. http://crypto.stanford. edu/pbc/.

作者简介：

杨旸（1984-），女，湖北随州人，博士，福州大学讲师、硕士生导师，主要研究方向为信息安全。

林柏钢（1953-），男，福建福州人，福州大学教授、博士生导师，主要研究方向为信息安全。

马懋德（1957-），男，天津人，南洋理工大学副教授、博士生导师，福州大学兼职教授，主要研究方向为信息安全。