基于属性的可搜索加密方案

李 双[1)] 徐茂智[2)]

[1)](北京工商大学理学院 北京 100048)

[2)](北京大学数学科学学院 北京 100871)

摘 要 2004年,Boneh等人利用匿名的基于身份的加密方案构造了一个公钥可搜索加密方案(PEKS),解决了特定环境下对加密数据进行检索的这一困难问题.已有的可搜索加密方案,通信模式往往是一对一的,关键词密文也只能被特定的单个用户查询或解密,这样的通讯模式在实际的系统中有很多局限性.作者首次给出了基于属性的可搜索方案(ATT-PEKS)的定义和构造算法.此方案与PEKS的不同之处在于基于属性的可搜索方案是适应群组的公钥加密搜索方案,扩大了信息的共享性,节省了第三方信息存储的空间.作者同时对此方案进行了一致性分析和安全性证明.

关键词 可搜索加密;基于属性的加密;双线性Diffie-Hellman问题;安全性证明;信息安全;网络安全

中图法分类号 TP309 **DOI号** 10.3724/SP.J.1016.2014.01017

Attribute-Based Public Encryption with Keyword Search

LI Shuang[1)] XU Mao-Zhi[2)]

[1)](*School of Science, Beijing Technology and Business University, Beijing* 100048)

[2)](*School of Mathematical Sciences, Peking University, Beijing* 100871)

Abstract In 2004, Boneh using anonymous hierarchical identity-based encryption scheme constructed a public key searchable encryption scheme (Public Key Encryption with Keyword Search shorthand for PEKS), which was proposed to solve the difficult task of the encrypted data to be retrieved under certain circumstances. Existing searchable encryption scheme, the mode of communication is often one-to-one; keyword cipher text can only be queried and decrypted by a particular individual user. There are a lot of limitations in this mode of communication in the actual system. We firstly present the definition of Attribute-Based Public Encryption with Keyword Search (ATT-PEKS) and the algorithm of construction. ATT-PEKS is different from PEKS, which based on the user's property is adapt to the group's public key encryption search program, expands the information sharing, saves storage space of third-party information. We also give the analysis of consistency and the proof of security.

Keywords PEKS; ABE; BDH; security proof; information security; network security

1 引 言

在网络迅猛发展、信息爆炸的今天,为了能快速查询到所需要的信息,搜索工具变得异常重要,如Google、百度已成为我们日常学习、生活的工具.如何使得搜索安全是一个重要的问题,于是一些关于加密数据的搜索技术应运而生.2004年,Boneh等

收稿日期:2012-10-29;最终修改稿收到日期:2013-12-23.本课题得到国家自然科学基金(10990011,61272499)、北京市属高等学校人才强教计划PHR(IHLB,201302)、北京市自然科学基金(1132002)、北京市哲学社会科学规划项目(13JGB036)资助.李 双,女,1977年生,博士,副教授,主要研究方向为信息安全、密码算法.E-mail:lishuang@th.btbu.edu.cn.徐茂智,男,1962年生,博士,教授,主要研究领域为信息安全、密码工程.

人[1]利用匿名的基于身份的加密方案构造了一个公钥可搜索加密方案(PEKS),该方案是针对在加密的邮件系统中邮件网关搜索带特定关键词邮件的应用场景而提出的.独立于文献[1]中的工作,Waters等人基于文献[2]提出了一个针对加密的审计日志进行检索的公钥加密方案[3].Park等人[4]提出了两个支持逻辑"与"的检索功能的可搜索公钥加密方案.2005年,Abdalla等人[5]对公钥可搜索加密方案的一致性进行了重新定义,将其划分为计算的、统计的和完美的三个类别.随后,他们给出了将匿名的基于身份的加密方案转化为具有计算一致性的可搜索公钥加密方案的通用方法,并基于Boyen和Waters[6]的匿名的基于层级身份的加密方案在标准模型中提出了一个基于身份的可搜索加密方案的一般定义,但没有给出具体的算法.2010年,Li等人[7]针对云计算环境提出了一个基于通配符的模糊的可搜索公钥加密方案.2011年,Cao等人[8]提出了一个可以同时对多个关键词进行检索的可搜索公钥加密方案.

已有可搜索加密方案大都是针对关键词密文与查询方之间的关系是一对一的情形,在越来越多的信息共享的大背景下,这些方案显然不适合关键词可被多方查询的需求.比如数字电视、视频点播、个人微博、个人藏品展示等都可以存储在如Google或Yahoo这样的第三方服务器上,此时数据与接收方是一对多的关系,也就是说数据可以被多个用户进行查询,传统的可搜索加密方案只能解决数据被唯一对应用户查询访问的问题.此时当加密方想要将一个信息和对应的关键词让目标的一群人得以分享,以现在的系统只能针对每一个人的私钥进行加密,将不同的机密结果在公共信道传送至对应的人群.查询者先是根据关键词查询到一些相关信息,然后再将对应消息解密.在这种情况下用已有的可搜索加密方案进行查询,无论是在效率上还是实际应用都是难以让人满意的.

基于属性的加密体制(ABE)[9],2005年首次被提出.在这种加密体制中加密者无需知道每个解密者的具体身份信息,而只需要掌握解密者一系列描述的属性,然后在加密过程中用属性定义访问结构对消息进行加密,当用户的密钥满足这个访问结构时就可以解密该密文.

因此我们利用基于属性的加密方案可以实现一对多的通信特性提出基于属性的可搜索加密方案,给出了一般定义、具体算法构造、复杂度分析和安全性分析.基于属性的可搜索加密方案与基于身份的可搜索加密方案相比,解除了关键词密文只能被唯一用户正确查询的限制,使得关键词密文可以被多个用户共享检索,节省了网络存储空间,提高了检索效率,特别适合网络迅猛发展下的各种应用.

2 基础知识

设 p,q 是素数,G_1 和 G_2 分别是阶为 p,q 的乘法循环群.称映射 $e:G_1\times G_1\rightarrow G_2$ 为一个双线性对,如果映射 e 满足下述性质[10]:

(1) 双线性(Bilinear)

对于任意 $a,b\in Z_p$ 和 $x,y\in G_1$,都有

$$e(x^a,y^b)=e(x,y)^{ab}.$$

对于任意 $x_1,x_2,y\in G_1$,有

$$e(x_1x_2,y)=e(x_1,y)e(x_2,y).$$

对于任意 $x,y_1,y_2\in G_1$,有

$$e(x,y_1y_2)=e(x,y_1)e(x,y_2).$$

(2) 非退化性(Non-degenerate)

存在 $x,y\in G_1$,使得 $e(x,y)\neq 1_{G_2}$.这里 1_{G_2} 代表 G_2 群的单位.

(3) 可计算性(Computable)

存在有效的多项式时间算法对任意的 $x,y\in G_1$,计算 $e(x,y)$ 的值.

在群 G_1 上,作为密码学的安全假设,有以下几个密码学困难问题.

定义1. 计算Diffie-Hellman问题(CDH):给定三元组 (g,g^a,g^b),其中 $a,b\in Z_p$ 且未知,g 为循环群 G_1 的一个生成元,计算 g^{ab} 的值.

定义2. 判定Diffie-Hellman问题(DDH):给定四元组 (g,g^a,g^b,g^c),其中 $a,b,c\in Z_p$ 且未知,g 为循环群 G_1 的一个生成元,判定 $c=ab(\mathrm{mod}\ p)$ 是否成立.

定义3. 双线性Diffie-Hellman问题[6](BDH):给定四元组 (g,g^a,g^b,g^c),其中 $a,b,c\in Z_p$ 且未知,计算 $e(g,g)^{abc}\in G_2$.

定义4. 判定双线性Diffie-Hellman问题(DBDH):给定五元组 (g,g^a,g^b,g^c,r),其中 $a,b,c\in Z_p$ 且未知,$r\in G_2$,g 为循环群 G_1 的一个生成元,判定是否 $r=e(g,g)^{abc}$.

目前仍未有解决BDH的有效算法,所以普遍认为BDH是一个困难问题.

定义5. 访问结构[10](Access Structure):一个实体集 $P=\{P_1,P_2,\cdots,P_n\}$,对于 $A\subseteq 2^P$ 是 2^P 上

的一个非空子集，$\forall B,C$，如果当 $B\in A$ 且 $B\subseteq C$ 时，则有 $C\in A$，那么称集合 $A\subseteq 2^P$ 是单调的. 一个访问结构 A（通常是指单调的访问结构）是 $P=\{P_1,P_2,\cdots,P_n\}$ 的一个非空子集合，即 $A\subseteq 2^P\setminus\{\Phi\}$. 包含于 A 中的集合称为授权集合（Authorized Set），而不包含在 A 中的集合称为非授权集合（Non-Authorized Set）.

定义 6.　Υ 表示一棵访问树，树中每个非叶子节点表示的是一个由子节点和阈值所描述的门限. 假设 num_x 表示节点 x 的子节点数目，k_x 表示节点 x 的阈值，则有 $0<k_x\leq num_x$. 当 $k_x=1$ 时，门限表示“或”门，而当 $k_x=num_x$，时，它表示“与”门. 另外，树中的叶子节点可以认为是一个属性且阈值 $k_x=1$.

为了简化对访问树的操作，定义了一些函数. 函数 $parent(x)$ 表示节点 x 的父节点. 如果 x 是叶子节点，则 $att(x)$ 表示与子节点 x 关联起来的属性值. 访问树 Υ 对每个节点的子节点进行编号，即子节点的编号是从 l 到 num，而函数 $index(x)$ 返回节点 x 的编号.

满足访问树：令 r 表示 Υ 的根节点，而 Υ_x 表示是访问树 Υ 中以 x 为根的子树，因此 Υ 也可以表示为 Υ_r. $\Upsilon_x(\gamma)=1$ 表示属性集合 γ 满足访问树 Υ_x. 我们可以通过以下方式递归去计算 $\Upsilon_x(\gamma)$ 的值. 如果 x 是一个非叶子节点，则计算 x 所有子节点 x' 的 $\Upsilon_{x'}(\gamma)$，当且仅当有至少 k_x 个子节点返回 1 时，$\Upsilon_x(\gamma)$ 返回 1；如果 x 是叶子节点并当且仅当 $att(x)\in\gamma$，则 $\Upsilon_x(\gamma)$ 返回 1.

3　基于属性可搜索加密方案的定义

3.1　ATT-PEKS 算法定义

首先给出基于属性的可搜索加密方案（Attribute Based PEKS，记为 ATT-PEKS）的定义. ATT-PEKS 是在基于属性的加密方案（ABE）的基础上提出来的，基于属性的可搜索加密方案引入属性的概念，从而使得满足相关属性的用户均可以对加密数据进行查询，不同于基于身份的可搜索加密，每个关键词密文与唯一查询用户对应.

定义 7.　一个基于属性的可搜索加密算法 ATT-PEKS 由 5 个多项式时间随机算法 Setup、KeyGen、PEKS、Trapdoor、Test 组成：

(1) $\mathrm{Setup}(1^\lambda)$：$(Pub,Msk)\leftarrow\mathrm{Setup}(1^\lambda)$

其中 λ 是系统参数. Setup 是由系统参数 λ 生成公共参数 Pub 和主密钥 Msk 的概率多项式算法；

(2) $\mathrm{KeyGen}(Msk,A)$：$Priv_A\leftarrow\mathrm{KeyGen}(Msk,A)$

密钥生成中心（KGC）根据访问结构 A 和系统主密钥 Msk 生成用户的私钥 $Priv_A$，其中 KeyGen 是由系统主密钥 Msk 和访问结构生成用户私钥的概率多项式算法；

(3) $\mathrm{PEKS}(Pub,w,\gamma)$：$C\leftarrow\mathrm{PEKS}(Pub,w,\gamma)$

发送方利用公共参数 Pub 和属性 γ 加密关键词 w，生成可搜索关键词的密文 C，只有属性满足访问结构的用户才能对关键词密文进行解密. 其中 PEKS 是由系统公共参数 Pub，关键词 w 和属性 γ 生成关键词密文的概率多项式算法；

(4) $\mathrm{Trapdoor}(Priv_A,w)$：$T_w\leftarrow\mathrm{Trapdoor}(Priv_A,w)$

接收方利用个人私钥 $Priv_A$ 计算查询关键词 $w\in\{0,1\}^*$ 的门限值 T_w 发送给网关服务器，其中 Trapdoor 是由用户私钥 $Priv_A$ 和关键词 w 生成关键词门限 T_w 的概率多项式算法；

(5) $\mathrm{Test}(Pub,T_w,C)$：$b\leftarrow\mathrm{Test}(Pub,T_w,C)$

已知公共参数 Pub，一个关键词 w' 的密文 $C=\mathrm{PEKS}(Pub,w',\gamma)$ 和关键词 w 的门限值 $T_w=\mathrm{Trapdoor}(Priv_A,w)$，如果 $w'=w$，输出 $b=1$，否则 $b=0$，其中 Test 是判断密文 C 是否是门限 T_w 对应关键词 w 的密文的概率多项式算法.

为了一致性，我们要求对所有的 $\lambda\in N$，所有的属性 S 和所有关键词 $w\in\{0,1\}^*$，有 $Pr[\mathrm{Test}(Pub,\mathrm{Trapdoor}(Priv_A,w),\mathrm{PEKS}(Pub,w,\gamma))=1]=1$. 这里的概率函数取遍所有公共参数 Pub，主密钥 Msk；取遍所有概率算法和所有随机预言机模型.

密钥管理中心由算法 KeyGen 根据访问结构 A 为用户生成私钥 $Priv_A$，然后用户由个人私钥 $Priv_A$ 和关键词 w 作为函数 Trapdoor 的输入生成关键词 w 的门限 T_w，并希望服务器根据门限 T_w 可以查找到关键词 w 的密文. 服务器用给定的门限 T_w 作为算法 Test 的输入来判断文件是否包含用户 Bob 指定的关键词 w 的密文. 且服务器不知道所查询关键词的明文 w，隐藏了查询信息的敏感性. 由于每个用户的密钥 $Priv_S$ 与属性有关，所以拥有属性相同密钥的用户可以对相同信息进行查询.

3.2　攻击游戏

KP-ATT-PEKS 的安全目标是在属性集合模型（Attribute-based Selective-Set Model）下达到选择关键词明文攻击的不可区分性安全. 以下是攻击者和挑战者之间在 Selective-Set 模型下的攻击游戏：

(1) 攻击者初始化：攻击者宣布它想挑战的属

性集 γ，并告知挑战者；

(2) 系统初始化：挑战者运行系统初始化算法，并将公共参数发送给攻击者；

(3) 第 1 阶段：攻击者可以选定一些访问结构 A_j 进行门限查询，其中对于所有的 j 要求 $\gamma \notin A_j$，攻击者可以选择关键词 w，询问其门限值 T_w；

(4) 挑战：攻击者提交两个关键词 w_0, w_1，然后挑战者随机选择其中之一 $w_b (b \in \{0,1\})$ 以属性集 γ 进行加密. 接着再将密文告诉攻击者；

(5) 第 2 阶段：此阶段是重复第一阶段的操作. 攻击者可以询问更多关键词的门限，限制条件是 $w \neq w_0, w_1$；

(6) 猜测：攻击者输出对 b 的猜测 $b' \in \{0,1\}$.

如果 $b'=b$，攻击者就获胜. 该攻击的优势定义为

$$Adv_{\text{ATT-PEKS}}^{\text{IND-CPA-SSM}} = |Pr[b'=b]-1/2|.$$

定义 8. 对于所有多项式时间的攻击者 A，如果 $Adv_{\text{ATT-PEKS}}^{\text{IND-CPA-SSM}}$ 都是可以忽略不计的，那么 KP-ATT-PEKS 在上述 Selective-Set 安全模型下是安全的.

4 算法构造

在 ATT-PEKS 方案中，我们的目标是实现查询方 Bob 发送一个关键词的门限值 T_w，这个门限值是利用 Bob 个人私钥和关键词 w 生成的. 网络服务器根据关键词门限值 T_w 搜索到所有包含关键词 w 的且属性满足访问结构的消息，服务器对关键词一无所知，但是服务器可以为满足属性的多个用户查询相同关键词的密文.

4.1 ATT-PEKS 算法构造

假设双线性 Diffie-Hellmen 问题(BDH)是难解的. 我们的构造是在 Goyal 等人[9]的基于属性的加密算法(ABE)基础上提出来的. 设双线性对 $e: G_1 \times G_1 \to G_2$，其中 G_1 是阶为 p 的乘法循环群，g 为群的一个生成元. G_2 是阶为 q 的乘法循环群，安全参数 k 确定群的大小. 同时定义拉格朗日系数 $\Delta_{i,s}(x) = \prod_{j \in s, j \neq i} \frac{x-j}{i-j}$，$i \in Z_p$，$S$ 是由 Z_p 中元素构成的集合，每个属性唯一对应 Z_p^* 中一个元素，$H_1: \{0,1\}^* \to G_1$，$H_2: G_2 \to G_2$，是两个 Hash 函数. 关键词在一个属性集 $|\gamma| = n$(由属性集 Z_p^* 中 n 个元素组成)下进行加密，ATT-IBEKS 由 5 个多项式时间算法构成：

算法 Setup(n)

输入：n，为属性集包含元素个数

输出：(Pub, msk)

1. $y, \alpha \xleftarrow{\$} Z_p$，$N=\{1,\cdots,n+1\}$
2. $t_1,\cdots,t_{n+1} \xleftarrow{\$} G_1$，$t_i \neq t_j$，$i,j \in N$
3. $g_0 \leftarrow g^{\alpha}$，$g_1 \leftarrow g^{y}$，$g_2 \xleftarrow{\$} G_1$
4. $T(X) \leftarrow g_2^{X^n} \prod_{i=1}^{n+1} t_i^{\Delta_{i,N}(X)}$
5. $Pub \leftarrow (g_0, g_1, g_2, t_1, \cdots, t_{n+1})$
6. $Msk \leftarrow (y, \alpha)$
7. Return(Pub, Msk)

算法描述：系统初始化(Setup)，首先选取两个随机数 $y, \alpha \in Z_p$，令 $g_0 = g^{\alpha}$，$g_1 = g^{y}$，$g_2 \in G_1$ 是随机选取的. 接着从 G_1 中随机选取互不相等的元素 $t_1,\cdots,t_{n+1}$，要求 $t_i \neq t_j$，对于 $i,j \in N$ 其中 $N=\{1,\cdots,n+1\}$.

同时定义一个函数 T，如下：

$$T(X) = g_2^{X^n} \prod_{i=1}^{n+1} t_i^{\Delta_{i,N}(X)}$$

T 可以看作函数 $g_2^{X^n} g^{h(X)}$，其中 h 是 n 阶多项式.

输出公共参数：$Pub = (g_0, g_1, g_2, t_1, \cdots, t_{n+1})$ 和主私钥：$Msk \leftarrow (y, \alpha)$.

算法 KeyGen(Pub, Msk, Υ)

输入：(Pub, Msk, Υ) 为公共参数 Pub，主密钥 Msk 以及访问树 Υ

输出：用户密钥 $Priv = (\alpha, \{Priv_x\}_{x \in \Upsilon})$

1. $r_x \xleftarrow{\$} Z_p$，$q_x(x) \xleftarrow{\$} Z_p^{k_x - 1}[x]$
2. $q_r(0) \leftarrow y$，$q_x(0) \leftarrow q_{parent(x)}(index(x))$
3. $D_x \leftarrow g_2^{q_x(0)} \cdot T(i)^{r_x}$，$i \leftarrow att(x)$
4. $R_x \leftarrow g^{r_x}$，$Priv_x \leftarrow (D_x, R_x)$
5. Return$(\alpha, \{Priv_x\}_{x \in T})$.

算法描述：密钥生成(KeyGen)，该算法以访问树 Υ，公共参数 Pub 和主私钥 Msk 作为输入，为用户(访问树中的节点 x)输出密钥. 持有这个密钥，用户可以解密由属性集 γ 加密的消息，只要满足条件 $\Upsilon(\gamma)=1$. 该算法执行的过程如下：

首先，为树中的每个节点 x 选择一个多项式 q_x，多项式的选择是从根节点 r 开始，以自上而下的方式进行选择的. 对于树中的每一个节点，多项式 q_x 的阶数 d_x 与这个节点的额阀值 k_x 存在以下关系：$d_x = k_x - 1$. 现在从根节点开始，对于根节点 r，令 $q_r(0) = y$，而多项式 q_r 在其他 d_r 个点的值完全进行随机选取. 往下的其他节点 x，令 $q_x(0) = q_{parent(x)}(index(x))$，而其他 d_x 个点的值随机定义，

其中函数 $parent(x)$ 表示访问树 Υ 中节点 x 的父节点.

经过上述操作所有多项式全部确定. 而对于每个叶子节点 x，将秘密信息 α，$Priv_x=(D_x,R_x)$ 发送给用户，其中 $D_x=g_2^{q_x(0)}\cdot T(i)^{r_x}$，$i=att(x)$；$R_x=g^{r_x}$，其中对于每个节点 x 对应的 r_x 都是从 Z_p 中随机选取的，其中函数 $att(x)$ 表示叶子节点 x 的属性；最后返回全部用户解密密钥 $Priv=(\alpha,\{Priv_x\}_{x\in\Upsilon})$.

算法 PEKS(Pub,w,γ)

输入：(Pub,w,γ)，分别为公共参数 Pub、关键词 w、和属性集 γ

输出：关键词密文 E

1. $s\xleftarrow{\$}Z_p$，$t\leftarrow e(H_1(w),g_0^s)$
2. $E\leftarrow(\gamma,E'=H_2(t)\cdot e(g_1,g_2)^s,$
   $E''=g^s,\{E_i=T(i)^s\}_{i\in\gamma})$
3. Return(E)

算法描述：加密关键词，假设要加密关键词 w，先计算其 Hash 函数 $H_1(w)$. 接着，选择一个随机值 $s\in Z_p$，计算 $t=e(H_1(w),g_0^s)$，然后计算关键词密文 E，并用属性集 γ 标记：

$$E=(\gamma,E'=H_2(t)\cdot e(g_1,g_2)^s,\ E''=g^s,\{E_i=T(i)^s\}_{i\in\gamma}).$$

算法 Trapdoor($Priv_x,w$)

输入：$(Priv_x,w)$，分别为节点 x 私钥 $Priv_x$ 和关键词 w

输出：关键词门限 T_w

1. $T_w=[W,U,V]\leftarrow(H_1^q(w),D_x,R_x)$
2. Return(T_w)

算法描述：生成门限，消息的接收方(Bob)由个人私钥 $Priv_x$ 和生成关键词 w 的门限值 $T_w=[W,U,V]$；其中 $W=H_1^q(w)$，$U=D_x$，$V=R_x$.

算法 Test(T_w,E,x)

输入：(T_w,E,x)，分别为关键词门限 T_w、关键词密文 E 以及节点 x

输出：判断值 b

1. $b\xleftarrow{\$}\{0,1\}$
2. $B\leftarrow\dfrac{E'}{DecryptNode(E,T_w,r)}$
3. If $H_2(e(T_w,E''))=B$
   then $b\leftarrow1$
   else
   $b\leftarrow0$
4. Return(b)

算法描述：判断，该算法以关键词的门限 T_w、关键词密文 E 和树的一个节点 x 作为输入，输出判断值 b. 现在有密文 E，访问树 Υ 以及属性 γ，如果 $\Upsilon(\gamma)=1$，则进行如下解密操作；否则返回“$\perp$”.

定义一个递归算法 DecryptNode(E,T_w,x)，它的输入是密文 E，门限 T_w 和树的一个节点 x，而输出群 G_2 上的一个值或者“$\perp$”. 具体操作如下：

令 $i=att(x)$，如果 x 是一个叶子节点，则有

$$\mathrm{DecryptNode}(E,T_w,x)=\begin{cases}\dfrac{e(D_x,E'')}{e(R_x,E_i)}, & i\in\gamma\\ \perp, & \text{其他}\end{cases},$$

其中

$$\frac{e(D_x\cdot g^r,E'')}{e(R_x\cdot g^r,E_i)}=\frac{e(g_2^{q_x(0)}\cdot T(i)^{r_x},g^s)}{e(g^{r_x},T(i)^s)}=e(g,g_2)^{s\cdot q_x(0)}.$$

如果 x 是非叶子节点，则考虑以下递归的情况. 算法 DecryptNode(E,T_w,x) 具体操作过程如下：对于节点 x 的所有子节点 z，令 $F_z=\mathrm{DecryptNode}(E,T_w,z)$. 设存在一个随机的大小为 k_x 的节点集合 S_x 且集合中的节点都是 x 的子节点，否则 $F_z=\perp$.

接下来，令

$$i=index(x),\ S_x'=\{index(z):z\in S_x\}.$$

计算：

$$\begin{aligned}F_x&=\prod_{z\in S_x}F_z^{\Delta_{i,S_x'}(0)}\\&=\prod_{z\in S_x}(e(g,g_2)^{s\cdot q_2(0)})^{\Delta_{i,S_x'}(0)}\\&=\prod_{z\in S_x}(e(g,g_2)^{s\cdot q_{parent(index(z))}})^{\Delta_{i,S_x'}(0)}\\&=\prod_{z\in S_x}e(g,g_2)^{s\cdot q_x(i)\cdot\Delta_{i,S_x'}(0)}\\&=e(g,g_2)^{s\cdot q_x(0)}.\end{aligned}$$

综上所述，现在已经得出了函数 DecryptNode 的完整定义，解密算法只需调用该函数在根节点的值. 所以，当且仅当密文属性满足访问树，即 $\Upsilon_r(\gamma)=1$. 则可计算出：$\mathrm{DecryptNode}(E,T_w,r)=e(g,g_2)^{y\cdot s}=e(g_1,g_2)^s$. 再由 $E'=H_2(t)\cdot e(g_1,g_2)^s$，所以判断算法根据 $E'/\mathrm{DecrptNode}(E,T_w,r)$ 是否等于 $H_2(e(T_w,E''))$ 的值来给出判断值 b.

4.2　ATT-PEKS 计算一致性

对于节点 x，当 $i=att(x)\in\gamma$，根据前面分析有

$$B=\frac{E'}{DecryptNode(E,T_w,r)}=\frac{H_2(t)e(g_1,g_2)^s}{e(g_1,g_2)^s}=H_2(t).$$

又由关键词门限 T_w 和 E'' 可计算

$$H_2(e(T_w,E''))=H_2(e(H_1(w)^a,g^s))=H_2(e(H_1(w),g^{as}))=B.$$

所以当 $H_2(e(T_w,E''))=B$ 说明关键词门限

T_w 与密文对应同一个关键词，且用户的属性满足 $i=att(x)\in\gamma$.

4.3 ATT-PEKS 复杂度分析

基本操作记为：P 表示双线性对，M 表示椭圆曲线群中的纯量乘法，E 代表指数运算，H 表示 Hash 函数. 假定关于访问树的运算预处理完成，我们用下面的运算量统计表给出具体运算量（设 $|G_1|=p$；属性集 γ 中的元素个数为 n，所有比特长度以 G_1 中一点的二进制表示的比特长为一个单位）. 其中 Crypt/PEKS 表示加密/计算关键词密文的算法；Decrypt/Test 表示解密/验证关键词门限值的算法；Trapdoor 表示计算关键词门限值的算法；Pub 表示公共参数（公钥）比特位数；Msg/w 表示消息/关键词的比特位数；ciph 表示密文的比特位数；Tw 表示关键词门限的比特位数. Algorithm 表示算法；Operator 表示操作；ABE 基于属性的加密方案[9]；ATT-PEKS 是基于属性的加密方案 4.1；PEKS 是可搜索加密方案[1].

表 1 ATT-PEKS 运算量统计表

Algorithm	Operator	ABE	ATT-PEKS	PEKS
Crypt/PEKS	P	1	2	1
	M	1	2	0
	E	$2+n$	$3+n$	2
	H	0	2	2
Decrypt/Test	P	0	1	1
	M	1	1	0
	E	0	0	0
	H	0	1	1
Trapdoor	P	0	0	0
	M	0	0	0
	E	0	1	1
	H	0	1	1
Size	Pub	$n+3$	$n+4$	$2g$
	M/w	1	任意	任意
	Ciph	$n+2$	$n+3$	$g+\log p$
	Tw	0	1	1

目前我们还没有看到关于基于属性的可搜索加密方案，所以就将 ATT-PEKS 与 ABE[9] 和第一个可搜索加密方案[1]进行比较，我们看出 ATT-PEKS 只是在计算关键词密文 PEKS 步骤比 PEKS[1] 计算量有所增加，其余步骤计算量基本一致；ATT-PEKS 与 ABE[9] 的计算量基本相当，说明在增加少量计算量的情况下实现了对加密关键词的多方可搜索功能.

5 安全性分析

下面我们来分析 ATT-PEKS 在基于属性的 Selective-Set 模型下的安全性，仍然采用规约化的思想，最终规约为求解 BDH 问题.

定理 1. 如果双线性 Diffie-Hellman（BDH）问题在群 G_1 上是难解的，则 ATT-PEKS 在基于属性的 Selective-Set Model 攻击下是安全的.

证明. 过程类似文献[9]. 假设存在一个多项式时间算法 A 在 Selective-Set 模型下以优势 ε_1 攻击 ATT-PEKS，我们构建一个算法 B 以 $\varepsilon\geq\frac{\varepsilon_1}{2}$ 的优势可以解 DBDH，以 $\varepsilon\geq\frac{\varepsilon_1}{eq_2q_T}$ 优势可以求解 BDH 问题.

设 $G_1=\langle g\rangle$，算法 B 是已知 g，$u_1=g^a$，$u_2=g^b$，$u_3=g^c\in G_1$，目标是输出 $v=e(g,g)^{abc}\in G_2$. 真正的挑战者随机选取 $\mu\in\{0,1\}$，输出 $\langle g^a,g^b,g^c,Z\rangle$，当 $\mu=1$ 时，$Z=e(g,g)^{abc}$，当 $\mu=0$ 时，随机选取 $Z\in G_2$. 算法 B 模拟挑战者和攻击者 A 之间的交互过程如下：

初始化：算法 B 调用 A，攻击者 A 选择属性集 γ，包含 Z_p 中 n 个元素的集合. 算法 B 设置公共参数 $g_1=g^a$，$g_2=g^b$，随机选取 n 次多项式 $f(x)$，计算 n 次多项式 $u(x)$ 满足：$u(x)=-x^n$，$\forall x\in\gamma$，并且 $u(x)\neq-x^n$，$\forall x\notin\gamma$. 因为 $u(x)$ 和 $-x^n$ 是两个 n 次多项式，所以它们至多有 n 个点相同，否则它们就是同一个多项式.

接下来算法 B 设 $t_i=g_2^{u(i)}g^{f(i)}$，其中 $i=1,2,\cdots,n+1$，因为 $f(x)$ 是 n 次多项式，t_i 是随机独立选取的，我们有 $T(i)=g_2^{i^n+u(i)}g^{f(i)}$.

阶段 1. 攻击者 A 询问一些属性集 γ 不满足的访问结构. 假设 A 请求访问结构满足 $\Upsilon(\gamma)=0$ 的密钥. 为了生成密钥，B 必须要为访问树中每个非叶子节点确定一个 d_x 次多项式 $Q_x(x)$.

$PolyUnsat(\Upsilon_x,\gamma,g^{\lambda_x})$ 是一个确定访问树中节点 x 的多项式（属性 γ 不满足以 x 为根节点访问树，即 $\Upsilon_x(\gamma)=0$）的过程，以访问树 Υ_x（x 为根节点），属性集 γ 和 $g^{\lambda_x}\in G_1$ 作为输入，其中 $\lambda_x\in Z_p$. 首先定义根节点 x 的 d_x 次多项式 q_x，满足 $q_x(0)=\lambda_x$. 因为 $\Upsilon_x(\gamma)=0$，x 有少于 d_x 个子节点满足属性，设 $h_x<d_x$ 是 x 的满足属性的子节点数. 对于每个满足属性 γ 的 x 的子节点 x'，随机选取 $\lambda_{x'}\in Z_p$，令 $q_x(index(x'))=\lambda_{x'}$，然后随机选取 q_x 剩余的 d_x-h_x 个点，至此 q_x 完全确定. 如此递归的定义访问树中不满足属性 γ 的点. 注意当已知 $g^{q_x(0)}$ 时 $g^{q_x(index(x'))}$ 可由插值法得到，并且有 $q_{x'}(0)=q_x(index(x'))$.

算法 B 调用 $PolyUnsat(T,\gamma,A)$ 来定义树中每个节点 x 的多项式 q_x，满足 $q_r(0)=\alpha$. 对于访问树 Υ 的每个叶子节点，如果 x 满足属性集合 γ，则完全知道 q_x，如果 x 不满足属性集合 γ，至少知道 $g^{q_x(0)}$.

算法 B 为每个节点 x 定义 $Q_x(\cdot)=q_x(\cdot)$，$y=Q_r(0)=b$. 每个节点 x 对应的密钥如下给出：这里 $i=att(x)$

If $i\in\gamma$, $D_x=g_2^{Q_x(0)}T(i)^{r_x}=g_2^{q_x(0)}T(i)^{r_x}$

$R_x=g^{r_x}$，其中随机选取 $r_x\in Z_p$

If $i\notin\gamma$, $g_3=g^{Q_x(0)}=g^{q_x(0)}$

$$D_x=g_3^{\frac{-f(i)}{i^n+u(i)}}(g_2^{i^n+u(i)}g^{f(i)})^{r'_x}$$

$$R_x=g_3^{\frac{-1}{i^n+u(i)}}g^{r'_x}$$，其中随机选取 $r'_x\in Z_p$

$i^n+u(i)$ 没有非零点，对于所有 $i\notin\gamma$，如果设 $r=r'_x-\frac{q_x(0)}{i+u(i)}$，那么

$$D_x=g_3^{\frac{-f(i)}{i^n+u(i)}}(g_2^{i^n+u(i)}g^{f(i)})^{r'_x}=g_2^{Q_x(0)}T(i)^{r_x}$$

$$R_x=g_3^{\frac{-1}{i^n+u(i)}}g^{r'_x}=g^{r_x}.$$

所以算法 B 可以构造具有访问结构 Υ 的密钥，并且分发与协议中所述相同.

接下来攻击者 A 询问随机预言机 H_1，H_2，询问关键词 w 的门限值 T_w 过程如同文献[1].

挑战：攻击者 A 选择两个关键词 w_0，w_1 请求给予挑战. 算法 B 随机选取 $b\in\{0,1\}$，给出关键词 w_b 的密文：

$C=(\gamma,E'=J\cdot Z,E''=g^c,\{E_i=(g^c)^{f(i)}=T(i)^c\}_{i\in\gamma})$，其中 Z 如前面定义.

阶段 2. 重复阶段 1，攻击者可以询问 w_0，w_1 以外关键词的门限.

猜测：算法 B 根据 A 的判断输出结果，A 给出猜测 b'，当 $b'=b$ 时，算法 B 输出判断 $c'=1$；当 $b'\neq b$ 时，输出 $c'=0$.

所以算法 B 求解判定性BDH问题的概率 $\varepsilon\geq\frac{\varepsilon_1}{2}$.

因为 $E''=g^c$. 意味着 $J=H_2(e(H_1(w_b),(g^a)^c))$，参照文献[1]中的分析对于随机预言机 H_2 查询和 H_2-list 列表中保留所有询问数对 (t,V).

其中 $t=e(H_1(w_b),(g^a)^c)=e(g,g)^{ac(b+a_b)}$，我们有 $\frac{t}{e(g^a,g^b)^{a_b}}=e(g,g)^{abc}$. 并且计算出 $e(g,g)^{abc}$ 的概率 $\varepsilon\geq\frac{\varepsilon_1}{eq_2q_T}$[1].

因此，设 H_1，H_2 是随机预言机，如果存在攻击算法可以攻击 ATT-PEKS，所占优势为 ε_1，那么存在攻击算法可以优势 $\varepsilon\geq\frac{\varepsilon_1}{2}$ 求解 G_1 中的 DBDH 问题；在 DBDH 可解的情况下，求解 G_1 中的 BDH 问题的概率 $\varepsilon\geq\frac{\varepsilon_1}{eq_2q_T}$，其中 q_1，q_2 是两个随机预言机的询问次数，q_T 门限的询问次数.　　证毕.

6　结　论

至此，我们首次给出了 ATT-PEKS 的算法构造并完成了安全性证明：如果存在攻击者可以攻击 ATT-PEKS，那么一定存在算法可以求解 BDH 问题. 因此 ATT-PEKS 在保证安全的前提下，提供了加密数据可被多方查询的功能，实现了加密数据的数据共享，节省了存储空间，提高了效率. 但 ATT-PEKS 方案还仅仅只是一个基础的原型方案，它主要是为了实现最基本的公钥加密的多方可搜索性. 随着研究的深入和实际应用的需要，该方案需要得到进一步的扩展.

参　考　文　献

[1] Boneh D, Crescenzo G D, Ostrovsky R, Persiano G. Public key encryption with keyword search//Proceedings of the EUROCRYPT'04. Interlaken, Switzerland, 2004: 506-522

[2] Boneh D, Franklin M. Identity-based encryption from the weil pairing. Advances in Cryptology-Crypto, 2001, 2139: 213-229

[3] Waters B R, Balfanz D, Durfee G, et al. Building an encrypted and searchable audit log//Proceedings of the Network and Distributed System Security Symposium 2004. San Diego, USA, 2004: 16-24

[4] Park D J, Kim K, Lee P J. Public key encryption with conjunctive field keyword search//Proceedings of the 5th International Workshop on Information Security Applications. Jeju Island, Korea, 2004: 73-86

[5] Abdalla M, Bellare M, Catalano D, et al. Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions. Journal of Cryptology, 2008, 21(3): 350-391

[6] Boyen X, Waters B. Anonymous hierarchical identity-based encryption (without random oracles)//Proceedings of the CRYPTO 2006. Santa Barbara, USA, 2006: 290-307

[7] Li J, Wang Q, Wang C, et al. Fuzzy keyword search over encrypted data in cloud computing//Proceedings of the 29th IEEE International Conference on Computer Communications. San Diego, USA, 2010: 1-5

[8] Cao N, Wang C, Li M, et al. Privacy-preserving multi-keyword ranked search over encrypted cloud data//Proceedings of the 30th IEEE International Conference on Computer Communications. Shanghai, China, 2011: 829-837

[9] Goyal V, Pandey O, Sahai A, et al. Attribute-based encryption for fine-grained access control of encrypted data//Proceedings of the 13th ACM Conference on Computer and Communications Security. New York, USA, 2006: 89-98

[10] Beimei A. Secure schemes for secret sharing and key distribution [Ph. D. dissertation]. Technion, Haifa, Israel: Israel Institute of Technology, 1996

LI Shuang, born in 1977, Ph. D., associate professor. Her research interests include information security, algorithm in cryptography.

XU Mao-Zhi, born in 1962, Ph. D., professor. His research interests include information security, cryptography engineering.

Background

In 2004, Boneh using anonymous hierarchical identity-based encryption scheme constructed a public key searchable encryption scheme (Public Key Encryption with Keyword Search shorthand for PEKS), which was proposed to solve the difficult task of the encrypted data to be retrieved under certain circumstances. Existing searchable encryption scheme, the mode of communication is often one-to-one; keyword cipher text can only be queried and decrypted by a particular individual user. There are a lot of limitations in this mode of communication in the actual system. We firstly present the definition of Attribute-Based Public Encryption with Keyword Search (ATT-PEKS) and the algorithm of construction. ATT-PEKS is different from PEKS, which based on the user's property is adapt to the group's public key encryption search program, expands the information sharing, saves storage space of third-party information. We also give the analysis of consistency and the proof of security.

This work is supported by National Natural Science Foundation of China (Grant No. 10990011, No. 61272499), Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (IHLB, 201302) and Beijing Municipal Natural Science Foundation (1132002).