抗内部关键词猜测攻击的高效公钥可搜索加密方案

王少辉 张彦轩 王化群 肖 甫 王汝传

(南京邮电大学计算机学院 南京 210003) (江苏省无线传感网高技术研究重点实验室 南京 210003)

摘 要 云环境下,如何对用户加密数据实现高效检索是学术界的研究热点。现有大部分公钥可搜索加密方案不能有效抵御由云服务器发起的内部关键词猜测攻击(Inside Keyword Guessing Attack,IKGA),而抗IKGA方案存在效率不高,以及相同关键词对应搜索陷门相同导致的关键词统计信息泄露等问题。鉴于此,提出了一个新的高效抗IKGA的公钥可搜索加密方案,并基于变形DLIN(Decision Linear Problem)假设,以随机预言机模型证明了新方案满足内部关键词猜测攻击下的语义安全。新方案中,搜索陷门包含随机数且相同关键词的搜索陷门不同。与其他PEKS方案相比,新方案减少了双线性对运算的使用次数,因此具有更大的性能优势。

关键词 可搜索加密,内部关键词猜测攻击,不可区分性

中图法分类号 TP309.7 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.07.014

Efficient Public-key Searchable Encryption Scheme Against Inside Keyword Guessing Attack

WANG Shao-hui ZHANG Yan-xuan WANG Hua-qun XIAO Fu WANG Ru-chuan

(College of Computer,Nanjing University of Posts and Telecommunications,Nanjing 210003,China)

(Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks,Nanjing 210003,China)

Abstract In the cloud environment,how to search users' encrypted data efficiently is the research hotspot in academic circle. Most current public-key searchable encryption schemes cannot effectively resist the Inside Keyword Guessing Attack (IKGA) launched by cloud servers,while the existing anti-IKGA schemes suffer the problems of low efficiency or the same search trapdoors generation algorithm for same keyword,which would reveal statistics information of keywords. This paper proposed a new efficient anti-IKGA public-key searchable encryption scheme,in which the search trapdoor is generated by a non-deterministic algorithm. Based on the modified DLIN (Decision Linear Problem) assumption,the new scheme is certified to satisfy semantic security against IKGA in the random oracle model. In the new scheme,the trapdoors are generated with random numbers thus same keyword has various trapdoors. Compared with other PEKS schemes,the new scheme reduces the number of bilinear pairing operations and thus has better performance advantages.

Keywords Searchable encryption,Inside keyword guessing attack,Indistinguishability

1 引言

随着云计算技术的迅猛发展,云端数据处理和存储服务不断升级,越来越多的用户选择将数据上传并存储至云端。为保证上传数据的隐私性和机密性,用户选择加密数据来防止信息泄露。然而,传统的基于明文关键词的检索模型不能解决密文下的关键词检索问题,因此可搜索加密概念被提出。用户首先将加密数据同关键词密文上传至服务器,拥有检索能力的用户可根据检索的关键词生成搜索陷门并发送给云服务器,云服务器则执行陷门搜索算法查找匹配密文文件,并将结果返回给用户用于解密。可搜索加密技术解决了不可信服务器下的数据传输与密文关键词检索问题,成为云安全领域的研究热点之一。

根据方案基于密码体制的不同,可搜索加密可分为对称可搜索加密和非对称可搜索加密。2000年,Song等[1]首次提出了对称密码机制下的高效可搜索加密方案,该方案通过线性扫描的方式查找关键词密文。2004年,Boneh等[2]针对邮件路由应用场景提出了非对称可搜索加密方案(Public key Encryption with Keyword Search,PEKS)。方案涉及数据发送方、接收方和云服务器。发送方使用接收方的公钥加密关键词和密文,并通过安全信道传输至云服务器。该方案保证服务器无法从关键词密文中获得任何关键词信息。文献[3]对PEKS进行了拓展,提出了多关键词公钥搜索加密方案。文献[4]提出了基于ElGamal同态加密的多关键词检索拓展方案,检索密文对外部攻击者不可见。Chang等[5]给出CP-ABE(Ciphertext Policy-Attribute Based Encryption)框架下的

到稿日期:2018-06-29 返修日期:2018-10-23

王少辉(1977-),男,博士,副教授,主要研究方向为密码学和信息安全,E-mail:wangshaohui@njupt.edu.cn(通信作者);**张彦轩**(1994-),男,硕士,主要研究方向为信息安全;**王化群**(1974-),男,博士,教授,主要研究方向为应用密码学、云计算安全和区块链技术;**肖 甫**(1980-),男,教授,主要研究方向为物联网和传感网;**王汝传**(1943-),男,教授,主要研究方向为信息安全和软件工程。

多用户可搜索加密方案，方案中密钥长度固定且独立于用户的属性，但该方案服务器端的关键词检索速度不佳。

理想状况下，关键词空间应为无限大。然而 Byun 等[6]指出，在实际应用中，用户通常会选取有具体含义的关键词，导致关键词空间的信息熵相对偏小。攻击者在获知关键词密文或搜索陷门后，可通过穷举常见关键词，实施离线关键词猜测攻击（Keyword Guessing Attack，KGA）以获得关键词信息。Jeong 等[7]证明了在关键词猜测攻击下不存在同时满足算法一致性和安全性的 PEKS 方案。为抵御关键词猜测攻击，Baek 等[8]在文献[2]的基础上提出指定服务器的 PEKS 方案（PEKS with designated server，dPEKS），在陷门搜索算法中引入云服务器的私钥，保证除指定的服务器外不存在能判断关键词密文和搜索陷门是否匹配的第三方。然而，Rhee 等[9]指出 Baek 等的方案仍不能抵御离线关键词猜测攻击，并于文献[10]中引入陷门不可区分性（Trapdoor Indistinguishability）的概念，证明了抗关键词猜测攻击的充分条件是陷门的不可区分性，并设计了一个基于该安全模型的高效 dPEKS 方案。Lu 等[11]在基于身份密码的系统下提出了满足密文不可区分性的 dIBEKS 方案，固定了检索陷门的变量个数以提高效率。

通常，关键词猜测攻击由外部敌手发起，但若 KGA 攻击由不可信云服务器自身发起，则用户隐私数据更易泄露。此种攻击方式被称为内部关键词猜测攻击（Inside KGA，IKGA）。而文献[1-11]中的方案均不能抵御 IKGA 攻击。为解决这一问题，Wang 等[12]引入双服务器模型，两个独立服务器通过分享秘密检索陷门执行搜索算法，但该方案需要服务器之间进行多次双线性对运算的交互，且只有在服务器非同谋时才能抵御 IKGA 攻击。文献[13]采用签密算法生成搜索密文，需使用额外的消息认证码供数据接收方验证返回结果，通信成本较高。最近，Huang 等[14]提出可搜索公钥认证加密方案（Public key Authenticated Encryption with Keyword Search，PAEKS），发送方使用自身私钥和接收方公钥生成搜索密文，接收方同样使用自身私钥和发送方公钥生成搜索陷门，服务器因此无法自行生成关键词陷门测试密文。然而该方案中的搜索陷门生成算法是确定性算法，关键词统计信息无法隐藏，因此仍存在信息泄露的问题。

本文提出了一个高效抗 IKGA 的公钥可搜索加密新方案，其中关键词密文和搜索陷门均采用非确定算法生成，避免了关键词统计信息的泄露问题。基于变形判断线性（Decision Linear Problem，DLIN）问题假设，我们在随机预言机模型下证明了新方案满足搜索陷门和关键词密文的不可区分性，因此新方案可提供内部关键词猜测攻击下的语义安全。与具有代表性的 PEKS 方案的性能相比，新方案具有更小的计算开销。本文第 2 节主要介绍所提方案所需要的预备知识，如双线性对、抗 IKGA 攻击的 PEKS 方案的形式化定义和安全需求；第 3 节给出抗 IKGA 的 PEKS 新方案的设计；第 4 节在随机预言模型下证明新方案满足抗 IKGA 的语义安全，并从性能方面与现有方案进行全面比较；最后总结全文。

2 预备知识

本节引入双线性对的概念和本文方案基于的困难性问题假设，并给出了抗内部关键词猜测攻击 PEKS 方案的形式化定义和安全定义。

2.1 双线性对与困难问题假设

首先给出可忽略函数的概念。

定义 1（可忽略函数） 考虑函数 $g(n)$，若对于任意多项式 $f(n)$，都存在正整数 N，当 $n>N$ 时，都有 $g(n)<1/f(n)$，则称 $g(n)$ 是可忽略的。

定义 2（双线性对） 设 G_1 和 G_T 是阶为大素数 p 的循环群，双线性运算 $e:G_1\times G_1\rightarrow G_T$ 满足以下条件：

1）双线性。对于 $\forall g_1,g_2\in G_1,a,b\in Z_p$，均有 $\hat{e}(g_1^a,g_2^b)=\hat{e}(g_1,g_2)^{ab}$ 成立。

2）非退化性。$\exists g_1\in G_1,g_2\in G_1$，满足 $\hat{e}(g_1,g_2)\neq 1_{G_T}$，其中 1_{G_T} 为群 G_T 中的单位元。

3）可计算性。对于 $\forall g_1,g_2\in G_1$，存在有效的计算方法计算 $\hat{e}(g_1,g_2)$。

Boneh 等[15]引入了判断线性困难问题假设，而 Huang 等提出的方案[14]则基于变形 DLIN（modified DLIN，mDLIN）困难问题假设。

定义 3（DLIN 困难问题假设） 给定 $\{g,g^x,g^y,g^{xr},g^{sy},g^z\}\in G_1$，其中 $x,y,r,s,z\in Z_p$ 均为随机数，g 为群 G_1 的生成元，DLIN 困难问题假设指对于任意的概率多项式时间算法 A，正确区分 g^{r+s} 与 g^z 的优势是可忽略的，即：

$$|\Pr[A(g,g^x,g^y,g^{xr},g^{sy},g^{r+s})=1]-\Pr[A(g,g^x,g^y,g^{xr},g^{sy},g^z)=1]|\leqslant negl(\lambda) \tag{1}$$

定义 4（mDLIN 困难问题假设） 给定 $\{g,g^x,g^y,g^{r/x},g^{sy},g^z\}\in G_1$，其中 g 为群 G_1 的生成元，$x,y,r,s,z\in Z_p$ 均为随机数，mDLIN 困难问题假设指对于任意的概率多项式时间算法 A，只能以可忽略的概率优势区分 g^{r+s} 与 g^z，即：

$$|\Pr[A(g,g^x,g^y,g^{r/x},g^{sy},g^{r+s})=1]-\Pr[A(g,g^x,g^y,g^{r/x},g^{sy},g^z)=1]|\leqslant negl(\lambda) \tag{2}$$

令定义 4 中 $x=y$，给出本文方案基于的困难问题性假设——mDLIN* 困难问题假设。

定义 5（mDLIN* 困难问题假设） 给定 $\{g,g^x,g^{r/x},g^{sx},g^z\}\in G_1$，其中 $x,r,s,z\in Z_p$ 均为随机数，g 为群 G_1 的生成元，mDLIN* 困难问题假设指对于任意的概率多项式时间算法 A，只能以可忽略的概率优势区分 $g^{r+1/s}$ 与 g^z，即：

$$|\Pr[A(g,g^x,g^{r/x},g^{sx},g^{r+1/s})=1]-\Pr[A(g,g^x,g^{r/x},g^{sx},g^z)=1]|\leqslant negl(\lambda) \tag{3}$$

2.2 PEKS 形式化定义

本节将给出 PEKS 方案的形式化定义和安全要求。如图 1 所示，方案涉及用户和云服务器两个实体。

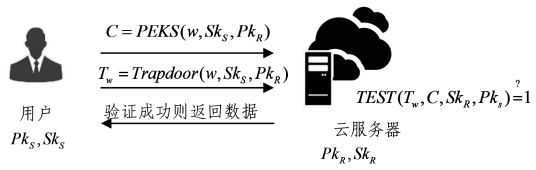

图 1 PEKS 系统框架

Fig. 1 PEKS system frame

用户自身的计算和存储资源有限，需要将数据文件及相关关键词加密后存储到云服务器，并在后续应用中利用关键

词陷门对加密文件进行检索。而云服务器通常拥有大量的存储空间和强大的计算能力，能为用户提供数据存储和密文检索服务，返回与陷门请求匹配的密文。

PEKS 方案通常由以下 5 个多项式时间算法组成。

1）$Setup(\lambda)$：输入系统安全参数 λ，输出系统公共参数 P。

2）$KeyGen(P)$：该算法由用户和服务器分别执行，以全局参数 P 为输入，用户输出公钥/私钥对（Pk_S，Sk_S），而服务器输出其公钥/私钥对（Pk_R，Sk_R）。

3）$PEKS(w,Sk_S,Pk_R)\rightarrow C$：该算法由用户执行，输入关键词 w、用户私钥 Sk_S 和服务器公钥 Pk_R，输出关键词密文 C，用户将数据密文和关键词密文一并存储在服务器中。

4）$Trapdoor(w,Sk_S,Pk_R)\rightarrow T_w$：当用户要查询关键词为 w 的数据时，其以 w、用户私钥 Sk_S 和服务器公钥 Pk_R 作为输入，输出为关键词查询陷门 T_w。

5）$Test(Sk_R,Pk,C,T_w)\rightarrow 0/1$：云服务器接收到关键词陷门 T_w 后，利用双方公钥 Pk 和自身私钥 Sk_R 对关键词密文 C 进行校验，如果关键词匹配，则输出 1，并返回相关数据密文；否则输出 0。

2.3　安全模型

在 PEKS 方案中，可能的安全威胁来自于云服务器和外部攻击者，这里只考虑内部云服务器。Rhee 在文献[10]中证明了抗关键词猜测攻击的充分条件是关键词密文和搜索陷门的不可区分性。因此，为抵御内部敌手关键词猜测攻击，方案需要保证敌手在已知服务器私钥的前提下满足：1）搜索陷门的不可区分性；2）关键词密文的不可区分性。我们通过多项式时间挑战者 Ch 与敌手 A 间的两个游戏 Game 1 和 Game 2，给出搜索陷门不可区分性和关键词密文不可区分性的定义。

Game 1　搜索陷门不可区分性

1）给定安全参数 λ，挑战者 Ch 生成全局参数 P、用户和服务器的公私钥对（Pk_S，Sk_S）和（Pk_R，Sk_R），并将 P，Pk_S 和服务器密钥对（Pk_R，Sk_R）发送给敌手 A。

2）敌手 A 可以适应性地询问如下预言机，访问次数通过多项式时间限定。关键词密文预言机 O_C：给定服务器的公钥 Pk_R 和关键词 w，O_C 预言机返回关键词密文 C 给敌手。搜索陷门预言机 O_T：给定服务器的公钥 Pk_R 和关键词 w，O_T 预言机返回搜索陷门 T_w 给敌手 A。

3）敌手 A 选择两个未询问过 O_C 预言机的关键词 w_0^* 和 w_1^* 并发送给挑战者。挑战者随机选择 $b\in\{0,1\}$，计算关键词 w_b^* 的对应陷门

$$T_{w_b^*}\leftarrow Trapdoor(w_b^*,Sk_S,Pk_R)$$

并返回给敌手 A。

4）敌手 A 可以继续访问 O_C 和 O_T 预言机，但要求不能就关键词 w_0^* 和 w_1^* 对关键词密文预言机 O_C 进行质询。

5）敌手 A 输出 $b'\in\{0,1\}$ 以猜测关键词 w_b^*，如果 $b'=b$，则敌手 A 获胜。

定义敌手 A 成功区分陷门的优势为：

$$Adv_A^T(\lambda)=|\Pr[b'=b]-0.5| \tag{4}$$

Game 2　密文的不可区分性

1）同 Game 1，挑战者 Ch 生成 P、公私钥对（Pk_S，Sk_S）和（Pk_R，Sk_R），并将 P，Pk_S 和服务器密钥对（Pk_R，Sk_R）发送给敌手 A。

2）同 Game 1，敌手 A 可以就关键词 w 对预言机 O_C 和 O_T 进行质询，质询次数通过多项式时间限定。

3）敌手 A 选择两个未询问过 O_T 预言机的关键词 w_0^* 和 w_1^* 并发送给挑战者。挑战者随机选择 $b\in\{0,1\}$，计算关键词 w_b^* 对应密文

$$C_b^*\leftarrow PEKS(w_b^*,Sk_S,Pk_R)$$

并返回给敌手 A。

4）敌手 A 可以继续访问 O_C 和 O_T 预言机，但要求不能就关键词 w_0^* 和 w_1^* 对搜索陷门预言机 O_T 进行质询。

5）敌手 A 输出 $b'\in\{0,1\}$ 以猜测关键词 w_b^*，如果 $b'=b$，则敌手 A 获胜。

定义敌手 A 成功区分关键词密文的优势为：

$$Adv_A^C(\lambda)=|\Pr[b'=b]-0.5| \tag{5}$$

定义 6　如果敌手 A 在多项式时间内 $Adv_A^T(\lambda)$ 和 $Adv_A^C(\lambda)$ 的优势是可以忽略的，则称该 PEKS 方案满足内部关键词猜测攻击安全性，即满足 IKGA 语义安全。

3　抗 IKGA 的高效 PEKS 方案

本节给出一个新的抗内部关键词猜测攻击的高效 PEKS 方案，其具体设计如下：

1）$Setup(\lambda)$：输入安全参数 λ 以生成公共参数 $\{p,g,h,G_1,G_T,\hat{e},H\}$，其中，$G_1$ 和 G_T 均为阶为 p 的循环群，g 和 h 为群 G_1 的随机生成元，$\hat{e}:G_1\times G_1\rightarrow G_T$ 是双线性映射，$H:\{0,1\}^*\rightarrow G_1$ 为散列函数。

2）$KeyGen(P)$：该算法由用户和服务器分别执行。用户随机选择 $x\leftarrow Z_p$ 作为其私钥 Sk_s，并计算公钥 $Pk_s=g^x$；而服务器同样选取随机数 $y\leftarrow Z_p$ 作为私钥 Sk_R，并输出其公钥 $Pk_R=g^y$，其中 Z_p 为质数集合。

3）$PEKS(w,Sk_S,Pk_R)\rightarrow C$：给定服务器公钥 Pk_R，用户随机选取 $r\leftarrow Z_p$，计算传输文件的关键词密文 $C=(C_1,C_2)$，其中：

$$C_1=H(w)^{Sk_S}\cdot(Pk_R)^{Sk_S\cdot r},C_2=h^r \tag{6}$$

最终，用户将密文 C 同加密数据一同上传至云服务器。

4）$Trapdoor(w,Sk_S,Pk_R)\rightarrow T_w$：用户选取随机数 $z\leftarrow Z_p$，利用其私钥 Sk_s 计算查询关键词 w 对应的搜索陷门 $T_w=(T_1,T_2)$。其中：

$$T_1=H(w)^{Sk_S}\cdot(Pk_R)^{Sk_S\cdot z},T_2=h^z \tag{7}$$

5）$Test(Pk,C,T_w)\rightarrow 0/1$：服务器接收到用户查询陷门 T_w 后，若关键词密文 C 满足 $\hat{e}(\frac{C_1}{T_1},h)=\hat{e}(Pk_s,\frac{C_2}{T_2})^{Sk_R}$，则输出 1 说明关键词匹配，并返回相应数据密文；否则输出 0。

方案正确性：给定查询关键词 w，用户公私钥对（Pk_S，Sk_S）=（g^x，x），服务器公私钥对（Pk_R，Sk_R）=（g^y，y），如果关键词密文和搜索陷门相匹配，则有：

$$\begin{aligned}\hat{e}(C_1/T_1,h)&=\hat{e}(H(w)^x g^{xyr}/H(w)^x g^{xyz},h)\\&=\hat{e}(g^{xy(r-z)},h)\\&=\hat{e}(g^x,h^{r-z})^y\\&=\hat{e}(g^x,h^r/h^z)^y=\hat{e}(Pk_s,C_2/T_2)^{Sk_R}\end{aligned} \tag{8}$$

显然可以通过 Test 算法的验证。

如果给定关键词与查询关键词不匹配，即 $w' \neq w$，由散列函数的抗碰撞性，有 $H(w) \neq H(w')$，此时：

$$\begin{aligned}\hat{e}(C_1/T_1,h) &= \hat{e}\left(\frac{H(w)^x g^{xyr}}{H(w')^x g^{xyz}},h\right)\\ &= \hat{e}\left(\frac{H(w)}{H(w')},h^x\right)\hat{e}(g^{xy(r-z)},h)\\ &\neq \hat{e}(g^{xy(r-z)},h) = \hat{e}(Pk_s,\frac{C_2}{T_2})^{Sk_R}\end{aligned} \tag{9}$$

服务器检验算法输出 0。因此，本文方案是正确的。

4 性能与安全性分析

本节在随机预言模型下给出新方案的安全性证明及其与现有代表性方案的性能比较。

4.1 安全性分析

通过引理 1 和引理 2 可以证明，基于 mDLIN* 困难问题假设，新方案在随机预言模型下满足关键词不可区分性和搜索陷门不可区分性，从而可以有效抵御内部关键词猜测攻击。

引理 1 假设 mDLIN* 问题假设是困难的，那么对于任意的概率多项式时间算法的敌手 A，能区分关键词密文的概率优势 $Adv_A^C(\lambda)$ 可忽略。

证明：假设存在概率多项式时间的敌手 A 能以不可忽略的概率优势 ε_A 正确区分关键词密文，我们构造新的概率多项式时间算法 B 以解决 mDLIN* 问题。即给定 mDLIN* 问题实例 $\{g, g^x, g^{r/x}, g^{sx}, g^z\} \in G_1$，算法 B 要判定 $Z = g^z$ 是否与 $g^{r+1/s}$ 相等。其中挑战者选取的随机数 b 用于确定 Z 值，当 b 值为 0 时，$z = r + 1/s$；b 值为 1 时，z 为循环群 G_T 中的任意元素。

B 首先随机选择 $t \in Z_p$，令 $h = (g^{sx})^t$，此时系统的公共参数为 $\{p, g, h, G_1, G_T, \hat{e}, H\}$；然后选择随机值 $y \in Z_p$，定义服务器的公私钥对 (Sk_R, Pk_R) 为 (y, g^y) 并发送给敌手 A，即云服务器知晓其对应私钥 y；用户的公钥设定为 $Pk_S = g^x$，即用户私钥 $Sk_S = x$。算法 B 需回答敌手 A 对如下预言机的质询。

哈希预言机 O_H：算法 B 维护一个列表 L_H 用于存放一个四元组 $\langle w_i, h_i, a_i, c_i \rangle$，其中 w_i 为关键词，列表 L_H 的初始值为空。就敌手 A 对关键词 w_i 的哈希质询，B 按如下方式应答。

1）如果关键词 w_i 已经在 L_H 中出现，算法 B 将返回 $H(w_i) = h_i \in G_1$。

2）否则，算法 B 以 $\Pr[c_i = 0] = \theta$ 的概率选取 $c_i \in \{0,1\}$。如果 $c_i = 0$，则算法 B 随机选择 $a_i \in Z_p$，并令 $h_i = g^{r/x} g^{a_i} \in G_1$；否则令 $h_i = g^{a_i} \in G_1$。

3）算法 B 将所得的四元组 $\langle w_i, h_i, a_i, c_i \rangle$ 存储于 L_H 中，并将 h_i 返回给敌手 A。

关键词密文预言机 O_C：敌手 A 就关键词 w_i 进行 O_C 预言机的质询时，算法 B 首先查询列表四元组 $\langle w_i, h_i, a_i, c_i \rangle$。如果 $c_i = 0$，算法 B 将终止操作，随机输出 b 的猜测 $b' \in \{0,1\}$；否则 B 随机选择 $r_i \leftarrow Z_p$，计算并返回敌手 A 的关键词密文 C_{w_i}：

$$C_{w_i} = (C_{w_i,1}, C_{w_i,2}) = PK_S^{a_i} \cdot (PK_S)^{y \cdot r_i}, h^{r_i} \tag{10}$$

搜索陷门预言机 O_T：敌手 A 就关键词 w_i 对 O_T 预言机提出质询时，算法 B 首先查询列表四元组 $\langle w_i, h_i, a_i, c_i \rangle$。如果 $c_i = 0$，算法 B 将终止操作，随机输出 b 的猜测 $b' \in \{0,1\}$；否则 B 随机选择 $\mu_i \leftarrow Z_p$，计算并返回敌手 A 的搜索陷门 T_{w_i}：

$$T_{w_i} = (T_{w_i,1}, T_{w_i,2}) = PK_S^{a_i} \cdot (PK_S)^{y \cdot \mu_i}, h^{\mu_i} \tag{11}$$

经过多项式次数的质询后，敌手 A 选择两个未询问过 O_T 预言机的挑战关键词 w_0^* 和 w_1^* 发送给算法 B。算法 B 首先查询列表 L_H 获知关键词 w_0^* 和 w_1^* 对应的四元组 $\langle w_0^*, h_0^*, a_0^*, c_0^* \rangle$，$\langle w_1^*, h_1^*, a_1^*, c_1^* \rangle$。若 $c_0^* = c_1^* = 1$，算法 B 终止此次挑战并随机输出 b 的猜测值 $b' \in \{0,1\}$；若 c_0^* 或 c_1^* 中有一个值为 0，不妨设 $c_{b'}^* = 0$，B 随机选择 $r' \in Z_p$，计算并返回关键词密文：

$$C^* = (C_1^*, C_2^*) = (Z \cdot (g^x)^{a_{b'}^*} \cdot (g^x)^{y \cdot r'}, g^{t/y} \cdot (h)^{r'}) \tag{12}$$

可以看出，当 $Z = g^{r+1/s}$ 时，有：

$$\begin{aligned}C_1^* &= g^{r+1/s} \cdot (g^x)^{a_{b'}^*} \cdot (g^x)^{y \cdot r'} = (g^{r/x} \cdot g^{a_{b'}^*})^x (g^{xy})^{(\frac{1}{sxy}+r')}\\ C_2^* &= h^{(\frac{1}{sxy}+r')} = (g^{sxt})^{(\frac{1}{sxy}+r')} = g^{t/y} \cdot (h)^{r'}\end{aligned} \tag{13}$$

其中，$\frac{1}{sxy} + r'$ 对敌手 A 满足随机分布。当 Z 是随机元素时，C_2^* 对 A 也是随机的。

敌手 A 可以继续访问 O_C 和 O_T 预言机，直至输出猜测结果 b''，但敌手 A 不能就关键词 w_0^* 和 w_1^* 对 O_T 预言机进行质询。如果 $b'' = b'$，算法 B 输出 $b' = 0$，即 $z = r + 1/s$；否则输出 $b' = 1$，即 z 为随机数。

假设敌手 A 进行了 q_H 次 O_H 预言机质询，q_T 次 O_T 预言机质询和 q_C 次 O_C 预言机质询，参照文献[14]，我们用 ter 表示算法 B 终止游戏的两种情况。

1）算法 B 模拟 O_T 或 O_C 预言机时，有 $c_i = 0$。由于 c_i 独立同分布，此时 B 终止游戏的概率 $\Pr[ter_1]$ 为：

$$\Pr[ter_1] = 1 - (1-\theta)^{q_T + q_C} \tag{14}$$

2）敌手 A 选择的挑战关键词中，$c_0^* = c_1^* = 1$，算法 B 终止游戏的概率 $\Pr[ter_2]$ 为：

$$\Pr[ter_2] = (1-\theta)^2 \tag{15}$$

推导得出算法 B 不会终止游戏的概率为：

$$\begin{aligned}\Pr[\overline{ter}] &= (1 - \Pr[ter_1])(1 - \Pr[ter_2])\\ &= (1-\theta)^{q_T+q_C}(1-(1-\theta)^2)\end{aligned} \tag{16}$$

同文献[14]，当 $\theta = 1 - \sqrt{\frac{q_T + q_C}{q_T + q_C + 2}}$ 时，算法 B 以不可忽略的概率优势 $\varepsilon_T = \frac{2}{e(q_T + q_C)}$ 终止游戏。假设 ε_{adv} 为敌手 A 攻破 PEKS 方案的优势，此时算法 B 通过敌手 A 攻破 mDLIN* 从而胜出的概率为：

$$\begin{aligned}\Pr[b' = b] &= \Pr[b' = b \wedge ter] + \Pr[b' = b \wedge \overline{ter}]\\ &= \Pr[b' = b | ter]\Pr[ter] + \Pr[b' = b | \overline{ter}]\\ &\quad \Pr[\overline{ter}]\\ &= \frac{1}{2}(1 - \Pr[\overline{ter}]) + (\varepsilon_{adv} + \frac{1}{2}) \cdot \Pr[\overline{ter}]\\ &= \frac{1}{2} + \varepsilon_{adv} \cdot \Pr[\overline{ter}]\end{aligned} \tag{17}$$

因此，$Adv_A^C(\lambda) = |\Pr[b = b] - 0.5| = \varepsilon_{adv} \cdot \Pr[\overline{ter}]$。由假设，$\varepsilon_{adv}$ 和 $\Pr[\overline{ter}]$ 均不可忽略，从而算法 B 以不可忽略的概率优势解决 mDLIN* 问题。这与 mDLIN* 问题的困难性假设矛盾，从而新方案满足关键词密文不可区分性。

可以看出，新方案中搜索陷门 T_w 与关键词密文 C 的构

造方式一致，因此由引理 1 可知，敌手能区分搜索陷门的概率优势可忽略，即有如下引理 2；同时综合引理 1 和引理 2，我们给出定理 1，可知新方案在 IKGA 攻击下是安全的。

引理 2　假设 mDLIN* 问题假设是困难的，那么任意概率多项式时间的敌手 A 能区分搜索陷门的概率优势 $Adv_A^T(\lambda)$ 可忽略。

定理 1　基于 mDLIN* 困难问题假设，本文的新 PEKS 方案在随机预言机模型下满足内部关键词猜测攻击下的语义安全。

4.2　性能比较

本节给出本文方案与其他 PEKS 方案的性能和安全性比较，如表 1 所列。性能比较主要针对实际运算中最消耗资源的双线性对运算 P、模幂运算 E 以及抗碰撞哈希函数运算 H；安全性比较则阐述方案是否具备抗 IKGA 攻击的能力，以及陷门算法是否是非确定性算法。

从性能角度看，本文方案的性能与不能抵御 IKGA 攻击的 Boneh 等[2]的方案相当；Shao[16]方案中 PEKS() 和 TEST() 要用到多个最耗时的双线性运算，性能明显低于本文新方案；而 Huang[14]方案共需 3 个双线性运算、4 个模幂运算，与本文的 2 个双线性运算、7 个模幂运算相当。从安全角度看，能抵御内部关键词猜测攻击的为文献[14]、文献[16]和本文提出的方案，但前两种搜索陷门算法均是确定性算法，容易造成关键词的统计信息泄露。综上所述，本文所提新方案在安全性和运算性能方面均具有较大优势。

表 1　新方案与其他方案的性能与安全性比较

Table 1　Performance and safety comparison of new scheme and other schemes

方案	PEKS()	Trapdoor()	Test()	确定性算法生成陷门	抵御 IKGA
Boneh 等[2]	2E+P+2H	E+H	H+P	是	否
Baek 等[8]	3E+2P+2H	E+H	P+H	否	否
Huang 等[14]	3E+H	E+P+H	2P	是	是
Shao 等[16]	9E+3P+3H	2E	5E+4P+H	是	是
本文	3E+H	3E+H	E+2P	否	是

结束语　在云计算环境下，如何对存储在云服务器的加密数据实现高效的密文检索是学术界和工业界研究的热点问题，而现有的可搜索公钥加密方案通常不能抵御来自云服务器的内部关键词猜测攻击。鉴于此，本文设计了一个新的 PEKS 方案，以非确定性算法生成搜索陷门，并基于陷门不可区分性和密文不可区分性的安全定义，在随机预言模型下证明了新方案满足抗 IKGA 语义安全。与其他 PEKS 方案相比，新方案使用了最少的双线性对运算，可提供更小的计算开销。

目前，新方案的安全性建立在随机预言模型下，如何设计标准模型下安全高效的 PEKS 方案亟须解决；本文方案的设计基于对称性双线性对，如何利用非对称双线性对进一步提高方案的效率仍需关注；除此之外，实现 PEKS 方案的多搜索关键词拓展也是我们下一步要研究的问题。

参 考 文 献

[1] SONG D X, WAGNER D, PERRIG A. Practical Techniques for Searches on Encrypted Data[C]// IEEE Symposium on Security & Privacy. Berkeley, CA, USA: IEEE Computer Society, 2000: 44-55.

[2] BONEH D, CRESCENZO G D, OSTROVSKY R, et al. Public Key Encryption with Keyword Search[C]// International Conference on the Theory and Applications of Cryptographic Techniques. Heidelberg, Berlin: Springer, 2004: 506-522.

[3] DONG J P, KIM K, LEE P J. Public Key Encryption with Conjunctive Field Keyword Search[C]// International Conference on Information Security Applications. Heidelberg, Berlin: Springer, 2004: 73-86.

[4] ZHANG R, XUE R, LIU L, et al. Oblivious Multi-Keyword Search for Secure Cloud Storage Service[C]// IEEE International Conference on Web Services. Honolulu, Hawaii, USA: IEEE Computer Society, 2017: 269-276.

[5] CHANG Y J, WU J L. Multi-user Searchable Encryption Scheme with Constant-Size Keys[C]// IEEE International Symposium on Cloud and Service Computing. Kanazawa, Japan: IEEE, 2018: 98-103.

[6] BYUN J, RHEE H, PARK H A, et al. Off-Line Keyword Guessing Attacks on Recent Keyword Search Schemes over Encrypted Data[J]. Lecture Notes in Computer Science, 2006, 4165: 75-83.

[7] JEONG I R, KWON J O, HONG D, et al. Constructing PEKS schemes secure against keyword guessing attacks is possible? [J]. Computer Communications, 2009, 32(2): 394-396.

[8] BAEK J, SAFAVINAINI R, SUSILO W. Public key encryption with keyword search revisited[C]// International conference on Computational Science and Its Applications. Heidelberg, Berlin: Springer, 2008: 1249-1259.

[9] RHEE H S, PARK J H, SUSILO W, et al. Trapdoor security in a searchable public-key encryption scheme with a designated tester[J]. Journal of Systems & Software, 2010, 83(5): 763-771.

[10] RHEE H S. Secure searchable public key encryption scheme against keyword guessing attacks[J]. Ieice Electronics Express, 2009, 6(5): 237-243.

[11] LU Y, WANG G, LI J, et al. Efficient designated server identity-based encryption with conjunctive keyword search[J]. Annals of Telecommunications, 2017, 72(5/6): 1-12.

[12] WANG C H, TU T Y. Keyword Search Encryption Scheme Resistant Against Keyword-Guessing Attack by the Untrusted Server[J]. Journal of Shanghai Jiaotong University (Science), 2014, 19(4): 440-442.

[13] SUN L, XU C, ZHANG M, et al. Secure searchable public key encryption against insider keyword guessing attacks from indistinguishability obfuscation[J]. Science China (Information Sciences), 2018, 61(3): 1-3.

[14] HUANG Q, LI H. An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks [J]. Information Sciences, 2017, 403-404: 1-14.

[15] BONEH D, BOVEN X, SHACHAM H. Short Group Signatures [C]// International Cryptology Conference. Heidelberg, Berlin: Springer, 2004: 41-55.

[16] SHAO Z Y, YANG B. On security against the server in designated tester public key encryption with keyword search[J]. Information Processing Letters, 2015, 115(12): 957-961.