针对公钥关键词搜索加密(PEKS)的关键词猜测攻击

——应用密码学结课报告

黄天波,刘永志,杨洲

2020 年 1 月 3 日

目录

摘要

报告从可搜索加密尤其是非对称可搜索加密的问题来源出发，介绍了可搜索加密的过程、非对称可搜索加密的定义，并详细阐述了针对非对称可搜索加密(PEKS)的关键词猜测攻击及其防御措施。

关键词：可搜索加密；非对称可搜索加密；PEKS；关键词猜测攻击

1 引言

可搜索加密问题源于文献 [1]：假设用户Alice试图将个人文件存放在一个诚实但具有好奇心的服务器，以降低本地资源开销。为保护文件隐私，须采用某种加密方式将文件加密后存储，使用传统分组密码，只有密钥拥有者才具备解密能力，意味着Alice在执行基于关键词的查询操作时，需要下载所有已上传的文件，完全解密后再检索，会带来两个问题：①如果Alice在服务器上已有大量文件，一一下载会占用大量的网络带宽，可能造成服务器堵塞；②对已下载的所有文件完全解密会占用大量本地计算资源，效率极低。

解决此类问题的加密技术称为可搜索加密(searchable encryption，简称SE)，该技术要求只有合法用户才具备基于关键词检索的能力，随着研究的推进，其应用不仅限于此：2004年，Boneh提出使用非对称可搜索加密(asymmetric searchable encryption,简称ASE)解决“不可信赖服务器路由问题”[2]，由此引出本篇报告重点关注的概念PEKS（公钥关键词搜索加密）。

为了解决“不可信赖服务器路由问题”，Boneh等人[2]最早提出PEKS(public key encryption with keyword search)概念，并基于BF-IBE[3]构造了第一个PEKS方案BDOP-PEKS，安全性可归结为BDH(bilinear Diffie-Hellman)数学假设。Khader[4]基于K-resilient IBE构造KR-PEKS方案，在标准模型下达到IND-CKA安全。Crescenzo等人[5]提出基于二次剩余中二次不可区分性问题(quadratic indistinguishability problem，简称QIP)的PEKS方案。Abdalla等人[6]针对PEKS算法一致性定义缺陷，提出统计一致性(statistically consistency)和计算一致性(computationally consistency)，并描述了从基于身份加密(identity-based encryption，简称IBE)到PEKS的一般变换算法IBE2PEKS。

文献 [7–9]指出了当前PEKS的一个较为严重的安全隐患：由于关键词空间远小于密钥空间，而且用户通常仅检索一些常用关键词，攻击者可借此实施关键词猜测攻击(keyword guessing attack，简称KGA)，进而证明了不存在满足算法一致性并且在KGA下是安全的PEKS方案。因此，抵御KGA意味着需对PEKS机制本身加以修改。

2 可搜索加密

2.1 可搜索加密过程

如图1所示，可搜索加密可分为4个子过程：

step1. 加密过程，用户使用密钥在本地对明文文件进行加密，并将其上传至服务器。

step2. 陷门生成过程。具备检索能力的用户，使用密钥生成待查询关键词的陷门，要求陷门不能泄露关键词的任何信息。

step3. 检索过程，服务器以关键词陷门为输入，执行检索算法，返回所有包含该陷门对应关键词的密文文件，要求服务器除了能知道密文文件是否包含某个特定关键词外，无法获得更多信息。

step4. 解密过程，用户使用密钥解密服务器返回的密文文件，获得查询结果。

图 1: 可搜索加密过程

2.2 两类基本问题

可搜索加密问题的提出，源于解决两类可搜索加密的基本问题：①不可信赖服务器的存储问题；②不可信赖服务器的路由问题。本文着重讨论与非对称可搜索加密有关的不可信赖服务器的路由问题。

不可信赖服务器路由问题 源于文献 [2]：Bob通过不可信赖邮件服务器向Alice发送包含某些关键词的邮件，要求服务器不能获取邮件内容和相关关键词信息，但需根据关键词将邮件路由至Alice的某个终端设备。例如，如果邮件的关键词为“urgent”，则服务器将邮件分配至Alice的手机，如果邮件的关键词为“lunch”，则服务器将邮件分配至Alice的电脑。

3 非对称可搜索加密

3.1 定义

3.1.1 算法描述

Boneh等人[2]在非对称密码体制中引入可搜索加密，提出PEKS(public key encryption with keyword search)概念，算法描述如下。

定义1(PEKS).非对称密码体制下可搜索加密算法可描述为PEKS=(KeyGen,Encrypt,Trapdoor,Test)

1) (pk,sk)=KeyGen(λ)：输入安全参数λ,输出公钥pk和私钥sk；

2) C_w=Encrypt(pk,W)：输入公钥pk和关键词W,输出关键词密文C_w；

3) T_w=Trapdoor(sk,W)：输入私钥sk和关键词W,输出陷门T_w；

4) b=Test(pk,C_w,$T_{w'}$)：输入公钥pk、陷门$T_{w'}$和关键词密文C_w,根据W与W'的匹配结果,输出判定值b∈{0,1}。

Boneh等人[2]提出不可信赖服务器路由问题的解决思路：Bob使用Alice的公钥pk加密邮件和相关关键词，并将形如(PKE.Encrypt(pk,MSG),PEKS.Encrypt(pk,W_1),⋯,PEKS.Encrypt(pk,W_n))的密文发送至邮件服务器。这里，PKE.Encrypt为公钥密码加密算法，MSG为邮件内容，W_1,⋯，W_n为与MSG关联的关键词。Alice 将T“urgent”或T“lunch”长驻服务器，新邮件到来时，服务器自动对其关联的关键词执行与T“urgent”或T“lunch”相关的Test算法，如果输出1，便将该邮件转发至Alice的手机或个人电脑。

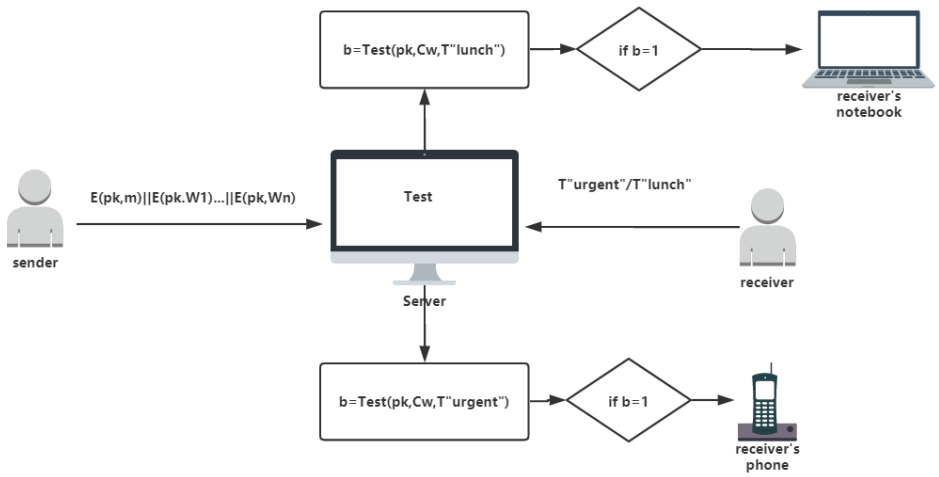

图 2: 非对称可搜索加密思路

3.1.2 算法一致性

加密算法的一致性是指解密与加密互为逆过程，即，对任意明文M，使用公钥pk加密后得到密文C，如果再使用pk对应的私钥sk解密，必能得到M。PEKS的一致性应满足：

①对任意关键词W，Pr[Test(pk,PEKS(pk,W),Trapdoor(sk,W))=1]=1；

②对任意关键词W_1、W_2且$W_1 \neq W_2$，Pr[Test(pk,PEKS(pk,W_1),Trapdoor(sk,W_2))]=1]=0。

文献 [2]中的方法无法满足上述要求[6]，鉴于此，Abdalla等人[6]对如上所述的完美一致性进行扩展，定义针对PEKS的计算一致性和统计一致性。

3.1.3 安全目标

PEKS需要满足：①没有陷门的服务器除文件长度外，无法获取任何文件信息；②拥有陷门T_w的服务器能够检索到所有包含W的密文文件。

3.2 关键词猜测攻击及防御措施

文献 [2]定义了关于PEKS的IND-CKA安全目标，并认为达到该目标的PEKS方案都是安全的。但是PEKS本身定义存在严重的安全隐患：文献 [7, 8]构造了针对PEKS方案[2, 10–12]的关键词猜测攻击，关键词猜测攻击是由于关键词空间远小于密钥空间，而且用户通常使用常用关键词进行检索，这就给攻击者提供了只需采用字典攻击就能达到目的的“捷径”。

3.2.1 攻击案例

Yau [8]等人构造了针对SCF-PEKS(Secure channel free PEKS,无需安全信道的PEKS)方案和PKE/PEKS方案的离线关键词猜测攻击，攻击者A按如下步骤执行离线关键词攻击：

①首先，A捕获一个有效的陷门T_w；

②A猜测一个合适的关键词w’，并计算出关键词密文$C_{w'}$；

③A根据接收者的公钥，$C_{w'}$和T_w执行Test算法，若返回值为1，猜测w’就是一个有效的关键词,否则，回到②。

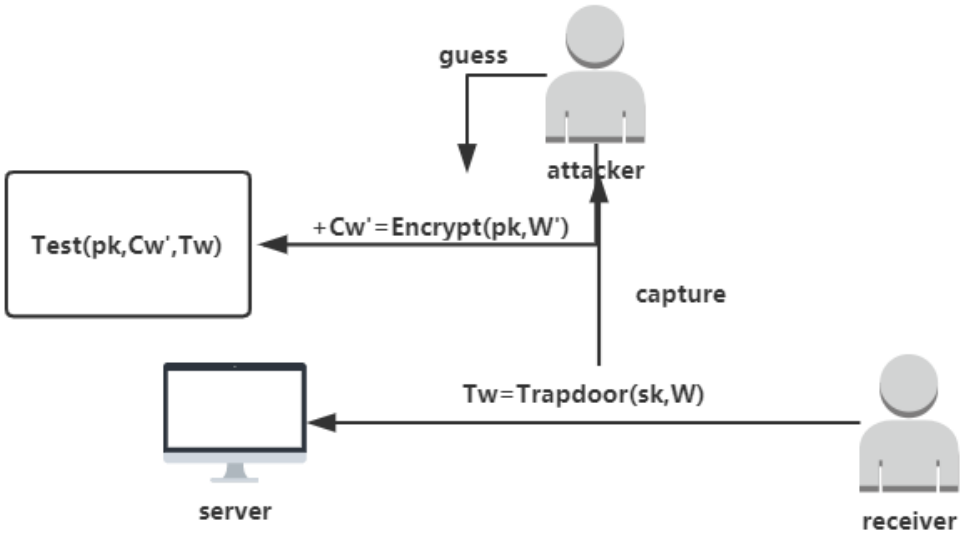

图 3: 关键词猜测攻击流程

进一步地，攻击者可以通过Test算法知道哪些PEKS密文包含该关键词。总地来说，执行离线关键词猜测攻击，攻击者需要先捕获一个陷门。如果在传送陷门的时候建立一个安全的信道，那么可怜的外部攻击者就束手无策了。SCF-PEKS方案给出了一个解决方法——服务器也持有一对公私钥对，发送者同时使用接收者和服务器的公钥加密明文，这样Test算法需要服务器的密钥作为输入，攻击者即使持有陷门也没有办法检测出包含特定关键词的邮件。当然，这个方法仍不能抵御内部攻击，因此，要设计能够抵御离线关键词猜测的PEKS方案，一个好的生成陷门的方法也非常重要。

导致关键词猜测攻击的原因可归结为:

(1) 关键词空间较小，且用户集中于使用常用词汇，给攻击者提供了遍历关键词空间的可能；

(2) PEKS算法一致性约束，使攻击者拥有对本次攻击是否成功的预先判定：执行Test算法，返回1说明本次攻击成功；否则，可以再继续猜测。

3.2.2 防御措施

1) 文献 [13]提出了PERKS方案，要求接收者在预处理过程中执行关键词注册算法，将输出的预标签通过安全信道传送给发送者，发送者才能为注册后的关键词W生成密文C_w。该方案实际上是对情形①的弥补，通过引入关键词注册过程，限制攻击者遍历关键词空间的能力。

2) 文献 [14]提出了PEFKS方案，在服务器端进行模糊陷门测试，过滤大部分不相关邮件，最后在本地精确匹配，得到检索结果。该方案通过引入模糊陷门，一定程度地降低了接收者外部PEKS算法的一致性，使其能够抵御关键词猜测攻击，但增加了客户服务器通信量和用户端计算量。

4 总结

PEKS能够适用于最基本的共享模式，具有广阔的应用空间。然而，其安全性问题是阻碍PEKS应用的重要原因。如前所述，几乎所有PEKS都遭受猜测关键词攻击的潜在威胁。虽然PERKS [13]和PEFKS [14]能够抵御关键词猜测攻击，但牺牲了一定的性能：PERKS [13]要求构建安全通道完成关键词注册；PEFKS [14]的返回结果包含非目标文件，需进行本地二次精确陷门测试。因此,设计一种更加安全、高效的PEKS扩展方案，是未来研究的方向之一。

参考文献

[1] Dawn Xiaodong, David Song, and Wagner Adrian Perrig. Practical techniques for searches on encrypted data.

[2] Christoph B?Sch, Pieter Hartel, Willem Jonker, and Andreas Peter. A survey of provably secure searchable encryption. *Acm Computing Surveys*, 47(2):1–51.

[3] Boneh Dan and Matt Franklin. Identity-based encryption from the weil pairing. In *Annual International Cryptology Conference*, 2001.

[4] Dalia Khader. Public key encryption with keyword search based on k-resilient ibe. In *Computational Science and Its Applications - ICCSA 2006, International Conference, Glasgow, UK, May 8-11, 2006, Proceedings, Part III*, 2006.

[5] G. Crescenzo and V. Saraswat. Public key encryption with searchable keywords based on jacobi symbols.

[6] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi. Searchable encryption revisited: Consistency properties, relation to anonymous ibe, and extensions. *Crypto*, 21(3):350–391.

[7] Wook Byun Jin, Hyun Suk Rhee, Hyun A Park, and Hoon Lee Dong. Off-line keyword guessing attacks on recent keyword search schemes over encrypted data. In *Secure Data Management, Third VLDB Workshop, SDM 2006, Seoul, Korea, September 10-11, 2006, Proceedings*, 2006.

[8] Wei Chuen Yau, Swee Huay Heng, and Bok Min Goi. Off-line keyword guessing attacks on recent public key encryption with keyword search schemes. In *Autonomic and Trusted Computing, 5th International Conference, ATC 2008, Oslo, Norway, June 23-25, 2008, Proceedings*, 2008.

[9] Ik Rae Jeong, Jeong Ok Kwon, Dowon Hong, and Dong Hoon Lee. Constructing peks schemes secure against keyword guessing attacks is possible? *Computer Communications*, 32(2):394–396.

[10] Joonsang Baek, Reihaneh Safavi-Naini, and Willy Susilo. On the integration of public key data encryption and public key encryption with keyword search. In *Information Security, 9th International Conference, ISC 2006, Samos Island, Greece, August 30 - September 2, 2006, Proceedings*, 2006.

[11] Joonsang Baek, Reihaneh Safavi-Naini, and Willy Susilo. Public key encryption with keyword search revisited. In *Proceeding sof the international conference on Computational Science and Its Applications, Part I*, 1970.

[12] Dj Park, Kihyun Kim, and Pj Lee. Public key encryption with conjunctive field keyword search. 2004.

[13] Tang Qiang and Liqun Chen. Public-key encryption with registered keyword search. In *Public Key Infrastructures, Services and Applications - 6th European Workshop, EuroPKI 2009, Pisa, Italy, September 10-11, 2009, Revised Selected Papers*, 2009.

[14] Peng Xu, Hai Jin, Qianhong Wu, and Wei Wang. Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack. *IEEE Transactions on Computers*, 62(11):2266–2277.