**在加密云数据上的能够验证的多关键字搜索**

目录

[一、 Introduction简介 3](#_Toc40102566)

[二、 Related work相关工作 4](#_Toc40102567)

[三、 Preliminaries预备知识 6](#_Toc40102568)

[四、 Problem formulations问题表述 7](#_Toc40102569)

[4.1. System and threat models系统和威胁模型 7](#_Toc40102570)

[4.2. Overview of VMKS scheme VMKS方案概述 8](#_Toc40102571)

[4.3. Security modle 安全模型 9](#_Toc40102572)

[五、 Construction of VMKS scheme VMKS方案的构建 10](#_Toc40102573)

[5.1. System initialization 系统初始化 10](#_Toc40102574)

[5.2. Key generation密钥生成 10](#_Toc40102575)

[5.3. Ciphertexts generation 密文的生成 10](#_Toc40102576)

[5.4. Trapdoor generation 10](#_Toc40102577)

[5.5. Ciphertexts retrieval 密文的检索 10](#_Toc40102578)

[5.6. Results verification 结果的确认 10](#_Toc40102579)

[六、 Analysis of VMKS scheme 分析VMKS方案 11](#_Toc40102580)

[6.1. Correctness 正确性 11](#_Toc40102581)

[6.2. Security analysis 安全分析 11](#_Toc40102582)

[6.3. Performance analysis 技术性能分析 11](#_Toc40102583)

[七、 Conclusions 结论 12](#_Toc40102584)

KeyWords:

Searchable encryption可搜索加密

Certificate management 证书管理

Key escrow密钥托管

Ciphertexts indistinguishability暗文不可分辨性

Signatures unforgeability 不可伪造的签名

Abstract

可搜索加密(SE)使用户能够搜索加密的数据，例如存储在远程云服务器中的数据。现有的基于证书、身份和属性的SE方案存在证书管理或密钥托管限制。此外，半诚实但好奇（semi-honest-but-curious cloud）的云可能进行部分搜索操作，并返回一部分搜索结果，以降低成本。在本文中，我们提出了一种基于密文的安全密码原语、可验证的多关键字搜索(VMKS：Verifiable Multiple Keywords Search)，它利用了基于身份的加密(IBE: Identity-Based Encryption )和无证书的签名技术。VMKS方案允许用户验证搜索结果的正确性，并避免证书管理或密钥托管限制。然后，我们将演示所提议的VMKS方案的安全性(即，该方案实现了密文的不可分辨性和签名的不可伪造性)。我们也使用真实世界的数据集来评估其可行性和效率。

# Introduction简介

在将数据(例如文本、图像、视频)外包给远程云服务提供商的过程中，数据所有者(个人和组织)通常会对其(敏感)数据进行加密，以确保数据机密性。此外，一些组织可能需要确保它们符合相关行业法规和隐私要求(例如新的欧盟通用数据保护法规)。

尽管在外包之前对数据进行加密有很多好处，但是搜索加密的数据(和数据集)仍然是一个很困难的事情。可加密搜索 (SE)的设计允许用户安全地搜索密文，基于预定义的关键字，并有选择地检索感兴趣的文件。SE方案的例子包括使用关键字搜索的公钥加密(public key encryption with keyword search, PEKS)，后者可以大致分为基于证书的关键字搜索方案和基于身份(或属性)的关键字搜索方案。在基于证书的关键字搜索方案中，数据所有者通过使用特定数据用户的公钥对其数据进行加密，从而共享其数据。密钥限制是证书管理，因为需要通过证书管理系统来验证证书和公钥。换句话说，可伸缩性在实践中可能是一个挑战。基于身份(或属性)的关键字搜索方案的密钥限制是密钥托管，因为受信任的中心可以解密系统中的任何密文。

许多研究人员试图解决这些限制。例如，在[42]中提出的关键字搜索方案是为了减轻大多数现有SE方案的限制，比如[12,13,16].。然而，在[42]中提出的关键字搜索方案假设云是诚实但好奇的，即云服务提供商将忠实地遵循已建立的协议，但同时，推断有价值的信息是很灵活的。在实际应用中，这样的假设通常是不够的，因为云可能返回不完整的搜索结果(例如，最小化计算和带宽资源)。因此，我们考虑一个半诚实但奇怪的云[2]，它执行所请求的搜索操作的一小部分，并在实践中返回不完整的搜索结果。然后，我们提供一个结果验证机制，通过向存储在云中的每个文件添加一个签名来确保搜索结果的准确性。我们还观察到，对于可验证的关键字搜索方案，有必要支持多关键字搜索，以最小化带宽资源和改善用户搜索体验(因为一个关键字搜索返回许多不相关的搜索结果)。

为了同时实现上述搜索功能，我们设计了一个密码学原语——以后称为可验证的多关键字搜索(VMKS)。后者允许通过利用现有的公共审计技术(如[33,34]中介绍的技术)，在加密(云)数据方案上执行搜索。也就是说，一个特定的数据用户可以通过表示索引结构和无证书签名来进行多关键词搜索，验证搜索结果的正确性。此外，VMKS方案可以减轻证书管理和密钥托管限制，并且具有固定的活动门和密文检索大小(在部署到resourcecon应变设备上时，这两个都是重要的特性)。我们将VMKS主要特点总结如下:

1. Multiple keywords search. 多关键字搜索。所提出的VMKS方案允许特定数据用户在不增加陷阱门尺寸和密文搜索大小的情况下，在单个搜索查询中发出多个关键字搜索，包括合取conjunctive关键字搜索和析取disconjunctive关键字搜索，提高了用户的搜索体验。（合取关键字搜索意味着每个结果包含所有查询关键字，而析取关键字搜索意味着每个结果至少包含查询关键字。在VMKS方案中，我们主要讨论了连接关键字的搜索。）
2. Search results verification. 验证搜索结果。VMKS方案允许通过在每个文件中附加一个签名来验证搜索结果的准确性。
3. Certificateless. 无证书的。为了消除现有SE方案中的证书管理和密钥托管限制，VMKS方案是无证书的。

本文的其余部分组织如下。第2节和第3节回顾了与拟议的VMKS计划相关的文献和背景。第4节：系统模型、威胁模型、方案定义和安全模型。第5节：VMKS方案的具体结构。第6节，演示VMKS方案的正确性，并对其安全性和性能进行评估。第7节是本文的结论。

# Related work相关工作

如前所述，已有大量关于SE的研究，文献中提出了不同类型的SE方案(如单关键字搜索[17]、多关键字搜索[14,15]和可验证的关键字搜索[21,31])证明了这一点。

首个对称设置的密文检索方案[9,11]由Song等人[29]提出，几年后Boneh等人[1]提出了首个非对称设置的PEKS方案[22,23]。从那时起，有大量的工作集中在扩展搜索功能和安全的SE和相关方案。例如，Kurosawa等人的[10]演示了如何以可验证的方式支持文档更新操作(即修改、删除和添加)，以及检测恶意服务器的欺骗行为。为了支持多关键字搜索和用户撤销，Miao等人构建了一个基本的加密原语，基于属性的关键字搜索分层数据方案。

传统的SE方案只支持精确的关键字搜索，这限制了系统的可用性，影响用户的搜索体验。因此，Li等人提出了一个模糊关键字搜索方案，利用编辑距离来处理微小的印刷错误和格式不一致。然而，在实践中，SE方案还应该支持多关键字(concontive关键字或disconcontive关键字)搜索，进一步缩小搜索范围，单个关键字搜索往往产生许多不相关的搜索结果[4,35,39]。Zhang等人提出了一种排序多关键字搜索方案，该方案支持多所有者模型中的结果相关性排序和动态密钥生成。这减少了返回所有搜索结果的需要。在一个更有挑战性但更现实的场景中，恶意云服务提供商可能出于自私的原因故意返回错误的搜索结果，比如最小化成本。因此，用户应该使用一些数据验证解决方案来确保搜索结果的正确性[8,28]。Sun等人提出了一种高效的基于树的索引结构，可以对返回的搜索结果进行真实性检查。Miao等人还利用多签名技术提出了共享多所有者环境下可验证的关键字连接搜索方案。

为了避免数据所有者为了验证数据用户的公钥而建立证书管理基础设施，Zheng等人利用基于属性的加密(ABE)[19,20]和bloom filter提出了可验证的基于属性的关键字搜索方案。出于类似的目的，Sun等人在多所有者设置下提出了基于属性的关键字搜索和高效用户撤销(ABKS-UR)方案。虽然两种方案都能验证搜索结果的正确性，但都不能解决密钥托管的局限性。Wang等人在他们的同态认证无证书签名(HA-CLS)方案中解决了证书管理和密钥托管的限制，但该方案不支持密文检索。郑等人[42]提出了一种无证书关键字搜索(certificateless keyword search, CLKS)方案，该方案实现了密文上的关键字搜索，但不保证搜索结果的真实性。

为了实现所有这些功能，我们提出了一种利用无证书签名方案和公共审计评估的可验证多关键字加密云数据方案[33,34]。与现有的SE方案不同，我们的VMKS方案避免了基于证书的关键字搜索方案固有的缺点，以及由于半诚实但好奇的云服务提供商导致的虚假搜索结果。另外，我们的VMKS方案支持多关键词搜索，不增加相应的活动门大小（trapdoor size）。（）我们还注意到，在本研究时，文献中还没有其他方案同时实现搜索结果验证、多关键字搜索、无证书和固定活板门大小——见表1。

在VMKS scheme中，trapdoor的大小不受查询关键字数量的影响，因为每个查询仍然在组G1中包含5个元素。

各种方案的功能比较。

“Function1”: Search result verification;

“Function2”: Multi-keyword search;

“Function3”: Certificateless;

“Function4”: Constant trapdoor size.

Table 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Schemes | Function1 | Function2 | Function3 | Function4 |
| VABKS [43] | √ |  |  |  |
| ABKS-UR [32] | √ | √ |  |  |
| CLKS [42] |  |  | √ | √ |
| VCKS [30] | √ | √ |  |  |
| HA-CLS [33] |  |  | √ |  |
| Re-dtPECK [39] |  | √ |  |  |
| VMKS | √ | √ | √ | √ |

# Preliminaries预备知识

在本节中，我们将回顾理解VMKS方案所需的相关背景资料。

令*x* ∈ *RX*表示从集合X中均匀随机选择的元素*x*, [1,Y]是整数集合{1,2,…, *ϒ* }, G1、G2为素阶p的两个乘性循环群，g为G1群的生成子群。e是双线性映射*G*1 × *G*1 → *G*2,，具有如下性质:

(1) Bilinearlity: *e(au, bv)* = *e(av, bu)* = *e(a, b)uv* for all *a, b* ∈ *RG*1, *u,v* ∈*R Z*∗*p*;

(2) Non-degeneracy: *e*(*g, g*) 1;

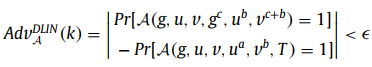
(3) Computablity: 对于*a, b* ∈ *RG*1，有一个计算 *e*(*a, b*) 的有效算法。

与我们的VMKS方案的安全性相关的安全假设描述如下:

Definition1：计算Diffie-Hellman (CDH)假设：设G1为p阶的可乘循环群，g是G1的生成元。给定元组 *g, ga, gb* ∈ *RG*1 和两个随机选择的元素*a, b* ∈*R Z*∗p，对于任何具有可忽略优势的概率时间对手*A*，*gab* ∈ *RG*1在计算上都是不可行的, 其中A的优势被定义为*Pr*[*ACDH(g, ga, gb)* = *gab*] *<* ().

Definition2：离散对数(DL)假设: 设G1是p的一组，g是G1的生成元。对于任何一个概率多项式时间对手A，它在解组G1中的DL问题上的优势是可以忽略的，定义为*Pr*[*A(g, ga)* = *a*] *< ,* where *a* ∈*R Z*∗*p*.

Definition3：决策线性(DLIN3)假设: 设G1是p的一组，g是G1的生成元。给定元组*(g, u,v, gc, ub,vc*+*b),* (*g, u, v, gc, vb, T*), 概率时间对手A的目标是区分组G1中的*va*+*b*和随机元素T，*u, v, T* ∈ *RG*1, *c, b* ∈*R Z*∗*p*. 那么我们说DLIN假设成立，如果A的优势*AdvDLIN A (k)*在打破DLIN问题中是可忽略的，其中优势*AdvDLIN A (k)* 由式(1)定义。

 ………………（1）

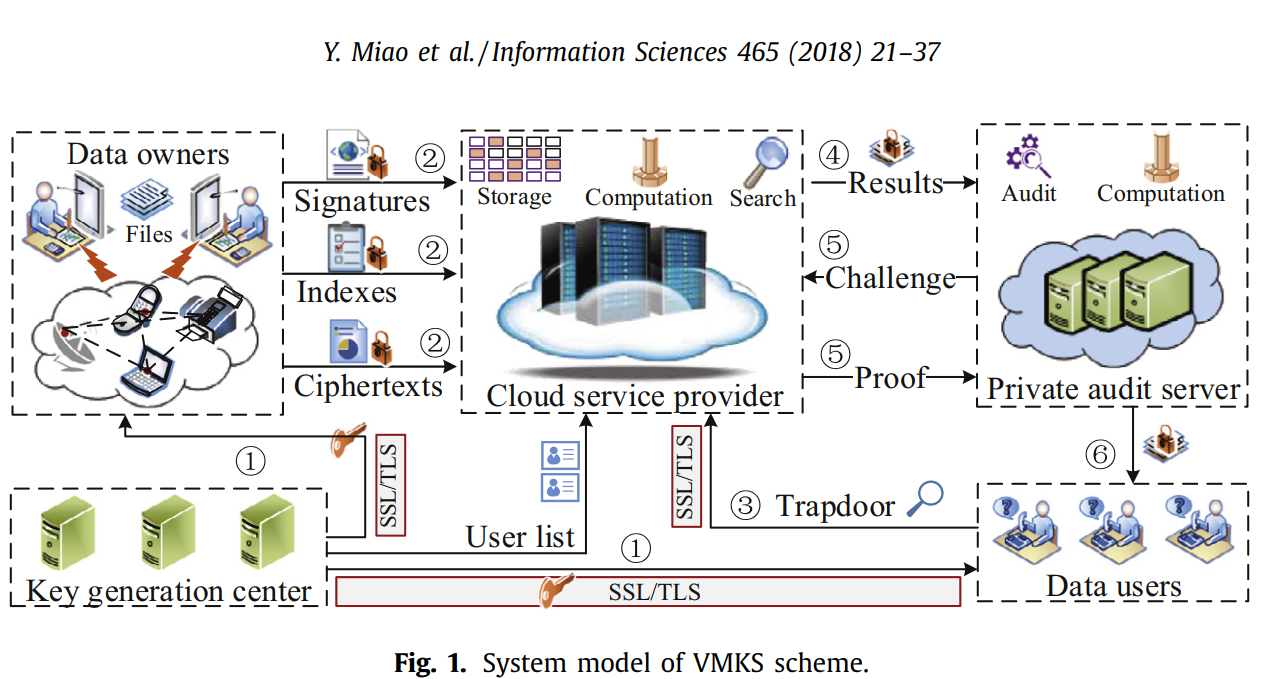
# Problem formulations问题表述

在本节中，我们给出了系统模型、威胁模型、方案定义和安全模型。

## 4.1. System and threat models系统和威胁模型

我们考虑的云存储系统包含五个实体，即: 数据所有者(DO)，多个数据用户(DUs)，

云服务提供商(CSP)、密钥生成中心(KGC)、私有审计服务器(PAS)，如图1（Fig. 1）所示。



Step1. 为了避免证书管理和密钥托管，KGC为DO和DUs生成部分密钥。

Step2. 其余的密钥将由各自的DO和DUs建立。

Step3. 然后，DO构建索引并将其外包给CSP，从而最小化本地存储和计算开销，实现密文检索。

Step4. 授权的DU可以通过向CSP提交一个活板门来进行多关键词搜索查询。

Step5. 一旦CSP检查了陷阱门与索引匹配，CSP就会将搜索结果返回给PAS，后者可以验证搜索结果的正确性。

Step6. 最后，将正确的搜索结果返回给DU。

每个实体的角色解释如下:

(1) KGC : KGC负责根据DO和DU的对应标识为它们生成部分私钥。私有密匙的其余部分由DO和DUs本身生成，如Step1所示。

(2) DO : DO收集数据文件并生成密文，包括文件密文、文件索引和文件签名。然后，这些生成的密文被外包给CSP -如Step2.

(3) DU : 当DU想要通过安全通道发出搜索查询时，例如SSL(安全套接字层)和TLS(传输)

他需要提交一个搜索查询(多个关键字)并生成一个陷阱门(或搜索令牌)，如Step3所示。

(4) CSP : 当CSP接收到活动门时，它用存储的索引匹配活动门，并将相关的搜索结果返回给PAS，如Step4所示。

(5) PAS : 为了验证搜索结果的正确性，PAS生成质询信息并发送给CSP, CSP需要返回证明,如Step5所示。然后，PAS检查证明信息是否能通过结果验证机制，如果搜索结果被认定不是伪造的，那么PAS就会把它们发送给DU;否则, PAS发送 ⊥ to DU, 如Step6所示.

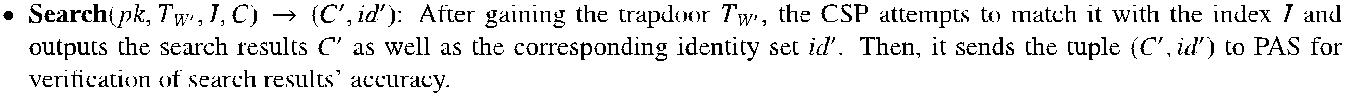
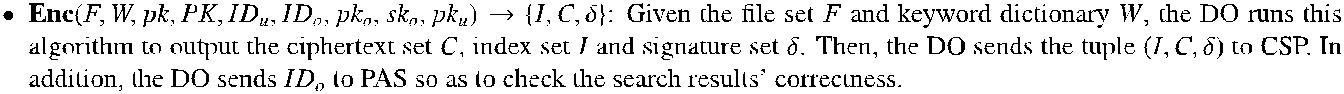
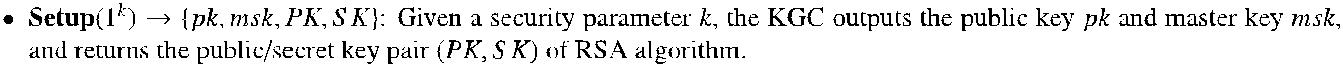
与之前的SE计划不同，我们不认为CSP是一个诚实但好奇的实体。具体地说，我们假设VMKS方案中的CSP是半诚实但好奇的。也就是说，CSP可能自私地执行部分搜索操作并返回部分搜索结果以最小化成本。我们还假设KGC是诚实但好奇的，类似于CLKS方案[42]中的假设。一个DU也可以模仿另一个DU，以获得有关目标DU的有价值的信息，但是模仿器不能与KGC串通。我们还假设PAS是完全可信的，但限制是PAS不能与CSP勾结。

## 4.2. Overview of VMKS scheme VMKS方案概述

VMKS方案是一组算法:Setup、ParKeyGen、KeyGen、Enc、Trap、Search和Verify， -见图2 (Fig.2)。

**Fig. 2.** Overview of the VMKS scheme.

|  |  |  |
| --- | --- | --- |
|  |  |  |



**Fig. 2.** Overview of the VMKS scheme.

## 4.3. Security modle 安全模型

与之前的无证书方案类似[6,42]，我们考虑两类对手，即:A1和A2。

• A1是一个外部攻击者，他被允许控制DUs的公钥，但主密钥msk除外。

• A2是一个内部攻击者，能够访问主密钥msk，而不控制DUs的公钥。

我们将通过下面两个安全游戏来说明VMKS方案对于这两个逆境都是安全的。让A1在第一场比赛中是对手，A2在第二场比赛中是对手，C是挑战者，保持以下两个列表:

TokenList: **：**对于元组(*IDu, w*)，与特定关键字w关联的搜索令牌已经被查询(A1, A2)。

UserInfoList: 给定元组 (*IDu, psku, sku, pku*, ϱ1, ϱ2, ϱ3)， 当Q1 = 1时，我们假设*psku*已经被(A1, A2)查询过, 当Q1 = 0时，它还没有被质疑. 当Q2 = 1, Q3 = 1时, 意味着*sku, pku*已经被*(A*1*, A*2*)*查询。否则，*sku, pku*还没有被这些对手质疑过。

具体来说，游戏1(或游戏2)是在A1(或A2)和C之间进行的。由于页面限制，我们将把讨论限制在游戏1，并将对游戏2感兴趣的读者(类似于游戏1，除了它没有阶段1)推荐到[6,42]。

Setup: C运行Setup来输出公共参数*pk*和主键msk，将pk发送到A1，并设置列表(TokenList和UserInfoList)为空。

Phase 1： 假设A1可以多项式地查询以下神谕，那么C可以进行以下算法:

(1) ParKeyGen:

(2) KeyGen:

(3) Replace-KeyGen:

(4) Trap:

Challenge phase:

Phase 2：

Guess:

Type-I adcersary.

Type-II adcersary.

# Construction of VMKS scheme VMKS方案的构建

基于无证书关键字搜索方案[42]，我们的目标是实现一个更实用的支持结果验证和多关键字搜索的SE方案。具体来说，VMKS方案应该验证搜索结果的有效性，并允许DU发出连接关键字搜索。对于无证书的结果验证，VMKS方案为每个文件附加一个签名，然后验证搜索结果的正确性。此外，该方案还需要避免证书管理和密钥托管。

## 5.1. System initialization 系统初始化

## 5.2. Key generation密钥生成

## 5.3. Ciphertexts generation 密文的生成

## 5.4. Trapdoor generation

## 5.5. Ciphertexts retrieval 密文的检索

## 5.6. Results verification 结果的确认

# Analysis of VMKS scheme 分析VMKS方案

## 6.1. Correctness 正确性

## 6.2. Security analysis 安全分析

## 6.3. Performance analysis 技术性能分析

#### 6.3.1. Theoretical performance 理论性能

#### 6.3.2. Actual performance 实际性能

# Conclusions 结论

随着越来越多的数据在服务和组织之间共享，搜索加密数据的能力将变得越来越重要。

在本文中，我们提出了一个可验证的多关键词搜索方案，该方案允许DUs在PAS的帮助下保证搜索结果的准确性。此外，该方案避免了证书管理和密钥托管，以及允许DUs在单个搜索查询中发出多个关键字。后者显著降低了计算和带宽成本。该方案还具有恒定的陷门和密文检索尺寸; 适合用于在资源受限的设备上进行部署。证明了VMKS方案实现了密文的不可分辨性和签名的不可伪造性，实验结果表明该方案是有效和实用的。

未来的研究包括扩展VMKS方案以支持动态操作，例如文件插入、修改和删除。