

对称可搜索加密的安全性研究进展

刘文心¹, 高莹²

¹北京航空航天大学 数学科学学院 北京 中国 100191

²北京航空航天大学 网络空间安全学院 北京 中国 100191

摘要 为节约本地存储空间以及管理开销,文件可通过云存储服务被上传到云服务器。云存储服务作为一项重要的云技术已得到了广泛的研究和应用。文件以明文的形式存储显然无法满足隐私保护和需求,但若以传统的加密方式将加密后的文件上传服务器又使服务器失去检索原文档的能力。可搜索加密(Searchable Encryption, SE)是近年来发展的一种支持用户在密文上进行关键字查找的密码学原语,它将用户的文件进行特殊的加密后上传到云服务器上,实现服务器可以根据关键字进行安全检索文件的功能,在方便用户使用的同时,也保护了文件的隐私安全。本文介绍了可搜索加密的基本概念,从对称可搜索加密的构建方法和加密手段出发,归纳总结了已有的对称可搜索加密的安全性结果。我们重点梳理了对称可搜索加密的适应性安全模型的发展历程,分析了推理攻击,文件注入攻击,以及新的安全模型与对抗手段,并指出目前可搜索加密安全性研究所面临的主要问题以及未来的发展方向。

关键词 云存储; 隐私泄露; 可搜索加密; 对称可搜索加密; 安全性

中图分类号 TP309 **DOI号** 10.19363/J.cnki.cn10-1380/tn.2021.03.05

A Survey on Security Development of Searchable Symmetric Encryption

LIU Wenxin¹, GAO Ying²

¹ School of Mathematics Sciences, Beihang University, Beijing 100191, China

² School of Cyber Science and Technology, Beihang University, Beijing 100191, China

Abstract To save local storage space and management overhead, now files can be uploaded to cloud servers. As an important cloud technology, cloud storage service has been widely researched and applied. The storage of files in plaintext obviously cannot meet the privacy and security requirements. However, if the encrypted files are uploaded to the server by traditional encryption, the server will lose the ability to search them by keywords. Searchable encryption (SE) is a cryptographic primitive developed in recent years that supports clients to perform keyword search on ciphertext. By searchable encryption, clients can encrypt the files and uploads them to the cloud server and then retrieve them by keywords, which is convenient for clients and protects the privacy of the files. This paper introduces the basic concepts of SE. Starting from the construction method of symmetric searchable encryption, we summarize the research results about security of existing symmetric searchable encryption. We focus on the development process of the adaptive security model of symmetric searchable encryption, and analyze inference attacks, file-injection attacks, new security models and countermeasures. Then we point out the main problems currently facing the security research of SE and the future direction.

Key words cloud storage; privacy leakage; searchable encryption; searchable symmetric encryption; security

1 引言

在当今的信息时代,云技术作为一种重要的技术得到了广泛关注和发展,云存储服务作为其较为核心的应用技术也备受人们的重视和研究。云存储服务可以为个人或者企业提供方便而又灵活的额外

存储空间而得到广泛应用。然而在大数据潮流下,数据隐私的泄露给人们带来巨大困扰,以明文存储的方式显然无法满足数据的隐私和安全需求。事实上,近年来由于黑客攻击,内部人员泄露或私下买卖用户数据库的事件层出不穷,保障云存储服务的数据隐私安全成为迫切需求。另一方面,如果用户将存储

通讯作者: 高莹, 博士, 副教授, Email: gaoying@buaa.edu.cn。

本课题得到基金项目支持: 北京市自然科学基金(No. M21033); 航天科学技术基金 (No. 2020-HT-BH-22); 国家自然科学基金 (No. 61932011, No. 61972017); 国家密码发展基金(No. MMJJ20180215)资助

收稿日期: 2020-04-17; 修改日期: 2020-05-18; 定稿日期: 2020-12-21

文件以传统加密的方式上传到云服务器又将导致另一个问题: 试想一个实际情形, Alice 希望将文件存储到云服务器, 出于隐私需求, Alice 可以通过传统加密方式预先加密好的文件并上传, 但只有 Alice 拥有密钥。当 Alice 希望查找含有某一关键字的文件时, 由于服务器对密文一无所知, Alice 需要从服务器上下载全部文件并一一解密以获得想要的文件, 这一方式显然效率极低。为了实现对加密文件的检索功能, 可搜索加密(Searchable Encryption, SE)技术应运而生。

可搜索加密技术是搜索技术和加密技术的结合。可搜索加密能够实现将用户的数据进行特殊的加密后上传到云服务器上, 并且可以实现根据关键字进行检索的功能, 有些可搜索加密方案更能实现范围查询或布尔查询等高级检索功能。在方便用户使用的过程中, 也保护了文件的隐私安全。

目前, 可搜索加密技术一般分为对称可搜索加密(Searchable Symmetric Encryption, SSE)和非对称可搜索加密(Asymmetric Searchable Encryption, ASE), 非对称可搜索加密目前一般又称为公钥可搜索加密(Public Key Encryption With Searching, PEKS), 下文在简称时统一用 PEKS 来指代。两者有不同的应用场景和构造方式。对称可搜索加密一般考虑单用户使用的情况, 相当于建立个人加密云盘, 依赖对称加密算法进行方案构造。而公钥可搜索加密一般考虑多用户使用的场景例如邮件系统或者多人文件共享系统, 主要依赖公钥加密算法进行构造。

近几年来可搜索加密的研究成为热点, 其发展非常迅速。已经有可搜索加密相关的综述性文章^[1-7]从整体上梳理了可搜索加密的研究和发展, 而对可搜索加密中安全性发展这条重要路线的梳理很少。其中, 文献[1-2]主要从搜索功能和应用场景对可搜索加密进行了梳理, 文献[3-4]比较具体的描述了之前典型的可搜索加密方案的方案细节以及做出对比, 同样比较详细地列举了可搜索加密的研究重点。文献[5-7]都对可搜索加密的模型进行了分类, 并从安全、效率、表达能力三个方面进行了总结, 粗略描述了各个方面的进展和挑战。不过以上综述性文章在安全性方面仍然缺少更多细节, 体现不了近年来研究中可搜索加密安全性研究的内在联系。因此, 本文从可搜索加密的安全性研究着手, 对其进行细致的整理和归纳总结, 对可搜索加密发展以来不断面临的新安全问题以及研究者们为此做的工作思路进行脉络整理。

根据以上可搜索加密综述性文章^[1-7]以及我们进

一步的总结发现, 对称可搜索加密的安全性主要考虑以下几个方面:

1. 上传文件本身的隐私。一般来说, 对称可搜索加密需要结合传统的对称加密手段如 AES 进行文件内容的加密, 但保证搜索功能是关键;
2. 用户搜索的关键字隐私。当用户搜索关键字时, 关键字内容与文件信息、用户习惯等隐私内容挂钩, 需要保护其安全性;
3. 加密数据库相关信息的隐私。相关信息如关键字频率, 搜索结果大小等容易被主动攻击的敌手利用并破坏其隐私性。

本文围绕以上三个方面对对称可搜索加密的安全性研究进行梳理。首先, 我们总结了已有的可搜索加密方案如何在保护文件内容的同时增加可搜索能力; 其次, 我们指出这些方案是否能够保证搜索时关键字的隐私, 以及如何保证隐私性; 最后, 面对近年来一些新型的利用加密数据库相关信息进行攻击的现象, 我们给出了对称可搜索加密的最新进展以及应对措施。

本文结构如下: 第 2 节简要介绍可搜索加密的基本概念与分类, 并着重对对称可搜索加密进行描述; 第 3 节梳理了对称可搜索加密安全性的相关研究成果, 并对其做出比较; 第 4 节具体地梳理有关对称可搜索加密的适应性安全研究进展与相关成果; 第 5 节是在第 4 节描述的适应性安全基础上, 对对称可搜索加密安全性的最新问题和研究方向进行了说明; 最后, 在第 6 节我们指出了目前可搜索加密面临的问题以及未来的发展方向。

2 可搜索加密的基本概念

2.1 非对称可搜索加密

2004 年, Boneh 等人^[8]将公钥加密应用到可搜索加密中, 基于 BDH(Bilinear Diffie-Hellman)困难问题假设, 首次提出了非对称可搜索加密(PEKS)的概念, 其主要目的是为了解决“多对一”的多用户邮件模型。例如 Bob, Carol 等多人向 Alice 发送邮件, 这些邮件通过 Alice 的公钥进行加密, Alice 可以根据私钥在邮件系统中进行关键字搜索并解密邮件内容并保证邮件服务器对内容一无所知。PEKS 算法描述如下:

定义 2.1. (PEKS). 一个 PEKS 系统由如下四个多项式时间的算法组成:

1. KeyGen(s): 输入安全参数 s , 生成 Alice 的一对公私钥对 A_{pub}, A_{pri} ;
2. PEKS(A_{pub}, W): 其他用户通过公钥 A_{pub} 和一个关键字 W 生成对 W 的加密;

3. $\text{Trapdoor}(A_{pri}, W)$: Alice 通过私钥 A_{pri} 和一个关键字 W , 生成陷门 T_W ;

4. $\text{Test}(A_{pub}, S, T_W)$: 给出公钥 A_{pub} , 一个加密 $S = \text{PEKS}(A_{pub}, W')$ 和陷门 $T_W = \text{Trapdoor}(A_{pri}, W)$, 验证是否 $W = W'$, 并输出结果 1 或 0。

实际上, Boneh 等人的 PEKS 方案可以视为在加密文件后附上加密的关键字清单, 用户可以通过私钥和 Test 算法与所检索关键字进行匹配, 以此进行关键字的搜索。随后 2006 年 Byun^[9]发现当前 PEKS 方案存在安全隐患: 由于关键字空间远小于密钥空间, 敌手可以通过离线关键字猜测攻击轻松破解 PEKS 体制。随后, 这一安全性问题由 Tang 等人^[10]提出注册关键字得以解决。目前公钥可搜索加密的安全性研究已经基本趋于成熟, PEKS 领域的研究主要集中于扩展方案的表达能力使其具备高级检索功能, 例如结合 ABE 实现访问权限的细粒度控制^[11-12]等。然而, 对称可搜索加密的安全性问题还未得到完全解决, 其安全性进展是目前研究的一个热点方向。因此, 本文接下来主要围绕对称可搜索加密的安全性进展进行梳理。

2.2 对称可搜索加密

2.2.1 SSE 的构建方法

2000 年, 对称可搜索加密由 Song 等人^[13]第一次提出, 用来解决“一对一”单用户进行云存储模型的方案, 即主要应用场景是个人用户将数据加密存储于云服务器, 并能在之后对加密数据进行检索。对称可搜索加密过程如图 1 所示。这之后, 可搜索加密迅速成为了被关注的热点研究问题。

SSE 的构建方法一般分为基于存储结构和基于索引两种方式。基于存储结构的 SSE 方案^[13-15]的构建方法一般通过特殊的加密手段将数据存储在特定的位置。例如 Song 等人^[12]的方案是应用流密码加密将密文在服务器内进行线性存储, 而 Naveed 等人^[16]的方案是利用伪随机函数和伪随机序列生成器将文件加密并“随机”分配到特定的随机位置, 只有拥有随机生成器“种子”的用户能准确找到文件的位置。然而, 基于存储结构构建 SSE 方案的效率往往比较低, 在搜索时需要服务器对整个存储器进行线性扫描并依次进行匹配, 而且对服务器的存储拓扑进行了严格要求。因此基于索引构建 SSE 方案是目前公认的主流方法, 绝大多数方案都是基于索引来构造的, 在本文接下来的讨论中我们默认讨论的是基于索引的 SSE 方案。基于索引的构建方法的优势在于不需要特定的加密手段和存储结构, 并且在搜索时有很高的效率, 其主要流程可以描述如下:

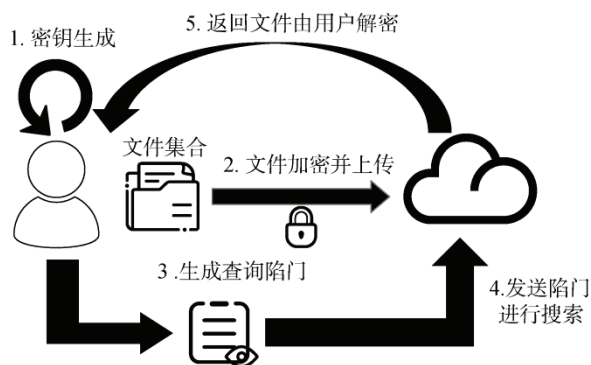


图 1 对称可搜索加密过程

Figure 1 Process of Symmetric Searchable Encryption

1. 用户选择安全通用的对称加密算法 SKE 例如 AES, 对数据文件加密, 保存密钥;
2. 用户根据文件内容构建索引结构, 保证索引内容与加密文件可以进行“链接”;
3. 将索引结构进行特殊加密, 和加密的数据文件一起发送给服务器;
4. 服务器存储加密索引和加密文件, 等待用户发送陷门进行操作。

如上所述, 这一类型的 SSE 构建方法中, 不同方案的核心在于第 2 步的索引结构的建立以及密文与索引进行“链接”的方式, 以及第 3 步对索引结构的特殊加密。而用户加密文件的手段使用的对称加密算法 SKE 和服务器存储文件的方式都是可选的, 具有灵活性。虽然基于索引的 SSE 相较于基于存储结构的 SSE 方案需要额外的索引构建过程, 以及额外存放索引的存储代价, 但是在效率上, 基于索引的 SSE 方案具有非常大的优势, 这在应用层面上决定了基于索引来构建 SSE 方案是最佳的。

2.2.2 SSE 的算法框架

在本文中我们参考 Curtmola 等人^[16]第一次正式定义的 SSE 的算法框架, 这一算法框架是基于索引的, 其定义如下:

定义 2.2.(SSE) 一个单用户的 SSE 方案的参与者包含一个用户和一个服务器, 假设 Δ 是关键字字典, $D \subseteq 2^\Delta$ 是文件集合, 用户希望将文件集合 D 存储与服务器上, 并且服务器可以提供对字典 Δ 的搜索服务, 一个基于索引的对称可搜索加密体制是指一个多项式时间算法的集合 $\text{SSE} = (\text{Gen}, \text{Enc}, \text{Trpdr}, \text{Search}, \text{Dec})$ 如下,

$K \leftarrow \text{Gen}(1^\lambda)$: 一个密钥生成的概率算法, 通过用户运行以建立系统, 以安全参数 λ 作为输入, 输出密钥 K 。

$(l, c) \leftarrow \text{Enc}(K, D)$: 由用户运行的一个概率

加密算法, 以密钥 K 和文件集合 $D = (D_1, \dots, D_n)$ 作为输入, 生成一个安全索引 I 和一系列密文 $c = (c_1, \dots, c_n)$ 。

$t \leftarrow \text{Trpdr}(K, w)$: 由用户运行的一个确定性算法, 根据希望检索的关键字 w 以及密钥 K 作为输入, 输出一个陷门 t 。

$X \leftarrow \text{Search}(I, t)$: 由服务器运行的一个确定性算法, 根据索引 I 和陷门 t 来查找文件集 D 中含有关键字 w 的文件, 并返回文件标识符集合 X 。

$D_i \leftarrow \text{Dec}(K, c_i)$: 由用户运行的一个确定性算法, 根据 X 中标识符得到对应密文, 用密钥 K 进行解密输出最终明文文件。

之后, Kamara^[17]在 Curtmola^[16]的方案上稍加改进, 增加了方案动态性以支持添加文件或删除文件的操作, 根据这两篇文章的工作, 可以把对称可搜索加密过程简化为归为以下 4 个步骤:

步骤 1. 建立和密钥生成过程: 用户对文件集合进行某种特殊加密后上传至服务器并生成密钥和加密数据库;

步骤 2. 陷门生成过程: 用户根据密钥和将要检索的内容生成特定陷门, 分为生成检索陷门和生成更新陷门, 并都上传给服务器;

步骤 3. 检索过程: 用户提交陷门, 由服务器根据陷门对加密数据库进行安全搜索和返回结果, 用户收到密文后解密得到最终结果;

步骤 4. 更新过程: 对于支持动态更新的可搜索加密, 可以通过将加密文件和更新陷门上传到服务器进行文件添加或删除操作, 注意添加操作和删除操作是区分开来的。

3 SSE 安全性的相关研究结果

在这一节, 我们首先粗略地梳理了对称可搜索加密安全性的相关研究成果, 并对其做出比较。之后在第 4 节我们将更加具体细致地分析它们的细节以及各个工作间内在的逻辑联系。

2000 年, Song 等人^[13]第一次提出 SSE, 他们方案的搜索时间与数据库大小成线性关系, 但在安全性和表达能力各方面都有欠缺。

2003 年, Goh 等人^[18]使用布隆过滤器首次建立了基于倒排索引的可搜索加密方案, 这篇文章中也正式地讨论了可搜索加密需要满足的安全模型, 但安全定义存在缺陷。2005 年, Chang 等人^[14]之后也提出了一种搜索时间为线性的方案, 并且在文章中同样提出了一个安全模型, 但是仍然存在缺陷。

2006 年, Curtmola 等人^[16]第一次正式定义了几种泄露信息, 指出了以往工作中安全模型存在的问题, 并在静态环境下设计了基于倒排索引的 SSE 方案, 实现了亚线性的搜索复杂度。

表 1 SSE 安全性相关工作的对比

方案	动态性	适应性安全	前向安全	效率
SWP00[13]	×	×	×	低
CM05[14]	✓	×	✓	低
CGKO06[16]	×	✓	×	中
KPR12[25]	✓	✓	×	中
SPS14[26]	✓	✓	✓	高
Bost16[20]	✓	✓	✓	高
Kim17[21]	✓	✓	✓	高

之后, 2012 年, Kamara 等人^[17]正式引入了动态对称可搜索加密(Dynamic Symmetric Searchable Encryption, DSSE)的概念, 动态性的引入要求 SSE 方案具有方便地添加或删除文件的能力。而之前的静态 SSE 方案中, 更新文件意味着重构索引, 代价高昂。这篇文章给出了实现亚线性搜索时间的 DSSE 方案, 但其在更新文件时会泄露更新关键字的哈希值。2014 年, Stefanov 等人^[19]提出前向安全的概念, 这是一个针对 DSSE 方案的安全概念, 具体含义我们将在第 4 节具体解释, Stefanov 指出这一安全概念在之前的 DSSE 方案中一直被忽略, 作者在文章中构建了一个动态的支持前向安全的亚线性搜索时间方案, 但存储结构比较复杂导致效率不高。之后, 前向安全的 DSSE 方案得到了广泛研究, 2016 年, Bost 等人^[20]给出了前向安全的正式定义, 并提出了一个前向安全的 DSSE 方案。2017 年, Kim 等人^[21]构建了支持高效更新的前向安全可搜索加密方案。

在安全性的研究中, 发现可能存在的信息泄露并完善已有的安全模型是普遍的研究方法。在这些工作中往往假设敌手是半诚实的被动敌手。在 Curtmola 较好地完善了 SSE 的安全模型后, 另一条工作路线则是对已有的 SSE 方案进行主动攻击的研究, 即恶意敌手模型。例如, 2012 年, Islam 等人^[22]针对 SSE 方案的信息泄露给出了分析并提出了攻击手段, 他们的工作随后被 Cash 等人^[23]进一步的总结并改善。2016 年, Zhang 等人^[24]提出的文件注入攻击说明了前向安全对于 SSE 的重要性。这些内容我们在第 5 节进行更详细的描述。

4 SSE 的适应性安全

如第3节所描述的, SSE 的安全性研究是不断进步和完善的, 而在这一节我们主要梳理的是 SSE 适应性安全模型的建立。

Song 等人^[13]提出首个对称可搜索加密方案, 即 SWP 方案, 见图2。其方法是: 将明文视为编码后的关键字流, 与密钥流进行异或后上传至服务器, 之后用户进行搜索需要用同样的密钥流依次解密并进行比对。显然这个方案在检索时需要服务器对密文依次全部扫描, 搜索时间与文件的大小成正比关系, 效率很低。

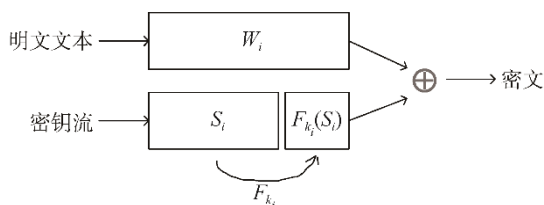


图2 Song 的 SWP 方案
Figure 2 Song's SWP scheme

SWP 方案在考虑安全性时将可证明安全理论中的 IND-CPA 安全(选择明文攻击下的不可区分性)引入, 保障了密文本身不会泄露隐私, 但还不足以描述可搜索加密面临的安全问题, 因为没有考虑到搜索过程中关键字的安全。2003 年, Goh^[18]提出了选择关键字攻击下的不可区分性即 IND-CKA 安全以及一个稍强的 IND2-CKA 安全, 两者区别在于是否为适应性敌手。IND-CKA 和 IND2-CKA 安全模型中, 敌手可以要求用户产生要求的密文文件以及查询陷门, 但无法做出和正常用户信息的区分并获得除此之外的更多信息。在此安全基础上, Goh 给出了基于布隆过滤器的具体方案, 相较于 Song, 这一方案引入了一份安全加密的倒排索引, 从而提高了搜索效率。2005 年, Chang 等人^[14]提出基于模拟的安全性定义, 考虑了敌手在实施攻击时即使能够获得之前所有轮次服务器端的查询结果情况, 但敌手在之后每一轮除了查询结果外, 无法获得任何信息。我们不详细描述 Goh 和 Chang 提出的安全模型, 是因为 4.2 节中随后 Curtmola 的工作指出了这两个安全模型的问题。

总而言之, 以上研究作为对称可搜索加密的初期工作, 在安全性的构建上虽有一些进展, 但缺乏正式安全模型定义, 或者定义中存在一些漏洞。

4.1 敌手模型与安全目标

在考虑 SSE 的安全性时, 我们首先需要考察敌手的威胁模型, 联系到现实背景的角度, 可搜索加

密方案通常以服务器作为敌手, 通常分为两类: 半诚实的敌手和主动攻击的敌手。半诚实的敌手会诚实地执行方案中的每一步协议, 但会尽可能解读协议过程中产生的信息并试图获取用户的隐私信息; 主动攻击的敌手往往会破坏正常的协议进程或进行额外的操作以此来威胁用户隐私。一般而言, 对 SSE 来说, 我们假设两种敌手拥有共同的能力包括:

- 可以观测到协议中用户与服务器间的全部交互
- 可以收集到协议进行过程中产生的信息
- 可以对文件的存储空间进行控制和访问

实际上, 现实场景中服务器作为敌手完全拥有这些能力, 因此我们所需确保的是在这些威胁之下方案仍能保护用户的隐私。故安全目标主要在于:

- 用户的数据本身具有隐私性, 敌手无法获取数据本身
- 用户的根据关键字查询时, 敌手无法获取关键字内容

4.2 SSE 的适应性安全模型

SSE 的适应性安全最初是由 Curtmola 等人^[16]给出定义, 之后由 Chase 等人^[25]对其进行了改进和简化。

4.2.1 Curtmola 的适应性安全模型

Curtmola 通过构造两个实例分别说明了 IND2-CKA 安全与 Chang 等人定义的安全性仍有漏洞: 这两个实例可以满足以上两个安全定义, 但是具有明显的漏洞可以让敌手还原用户的搜索内容。Curtmola 给出了一个新正式定义的安全模型, 构造了一个基于索引的对称可搜索加密方案。我们首先给出 Curtmola 的安全模型定义如下:

定义 4.1.(q-查询历史). 假设 Δ 是关键字字典, $D \subseteq 2^A$ 是文件集合, 一个在 D 上的 q -询问历史是一个包含了文件集合 D , 以及一个 q 个关键字的向量 $w = (w_1, \dots, w_q)$ 的双元组 $H = (D, w)$ 。

定义 4.2.(访问模式). 假设 Δ 是关键字字典, $D \subseteq 2^A$ 是文件集合, 对于一个在 D 上的 q -询问历史 $H = (D, w)$ 来说, 访问模式是指一组向量 $\alpha(H) = (D(w_1), \dots, D(w_q))$ 。

定义 4.3.(搜索模式). 假设 Δ 是关键字字典, $D \subseteq 2^A$ 是文件集合, 对于一个在 D 上的 q -询问历史 $H = (D, w)$ 来说, 搜索模式是指一个对称二进制矩阵 $\sigma(H)$, 其中对 $1 \leq i, j \leq q$, 第 i 行, 第 j 列的元素为 1 若 $w_i = w_j$, 否则为 0。

定义 4.4.(视图). 敌手关于查询历史 H 的视图 $V_K(H) = (I, C, T_1, \dots, T_q, id(D_1), \dots, id(D_n))$ 包括密钥 K 产生的加密文件及其索引, 历史查询陷门, 返回文

件标识符信息。

对以上定义, 访问模式表示方案中对存储系统进行访问时返回结果中发生的信息泄露, 比如搜索后返回结果的大小等。搜索模式则表示用户进行搜索时所泄露的信息, 比如两次询问是否是重复的。而视图实际上指在 q 次询问中服务器所掌握的全部已知信息。

Curtmola 分别在自适应(adaptive)和非自适应(nonadaptive)模型下形式化地定义了 SSE 的语义安全(semantic security, 简称 SS 安全)和不可区分性安全(indistinguishability security, 简称 IND 安全), 从而定义了四种模型。其安全性定义和密码学中可证明安全的定义类似, 都是通过敌手和模拟机 S 间进行

交互实验, 具体交互如图 3 所示^[2]。

实际上, 自适应(非自适应)指的是服务器在 q 次查询中能(否)根据返回信息进行查询调整的能力, 而语义安全是指加密文件、加密索引和陷门信息这些包含在一次 q -查询历史 H 中的内容的隐私性, 而不可区分性安全则是考虑用户不同的 q -查询历史 H_1, H_2 对服务器来说不可区分。另外 Curtmola 还说明了它们具有如下关系:

- 非自适应模型下: SS 安全与 IND 安全相互等价
- 自适应模型下: SS 安全包含 IND 安全, 即有了 SS 安全就有 IND 安全

因此, Curtmola 所定义的 SS 安全比 IND 安全具备更高的安全度。

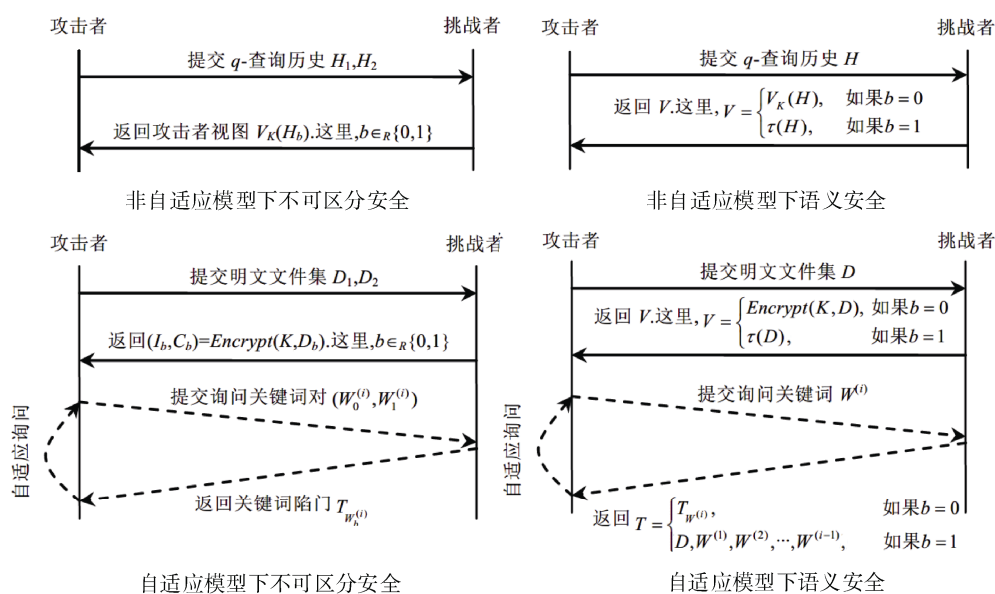


图 3 Curtmola 定义的安全模型

Figure 3 Security model defined by Curtmola

4.2.2 Chase 等人优化的适应性安全模型

实际上, Curtmola 所定义的四中安全模型中, 自适应性模型下的 SS 安全是安全性最强的, 但它仅仅是针对静态 SSE 方案。另外, 其中所定义的查询历史, 访问模式, 搜索模式, 视图等概念在表达不同 SSE 方案的泄露信息时具有局限性, 不能完全代表不同方案中实际泄露的全部信息。而为了更一般的对不同方案中产生的泄露信息进行说明, Chase 等人^[25]在他们的安全模型中提出了更为抽象的一个概念: 泄露函数。具体来说, Chase 用泄露函数来统一表达 SSE 方案在建立过程和检索过程中泄露的全部信息。根据 SSE 方案的步骤, 这些泄露信息被划分为两个部分: 建立过程中的泄露函数 L_{Setup} 和检索操作过程中的泄露函数 L_{Search} 。这样做的好处在于, 不同 SSE

方案的构建者可以不用繁琐的定义查询历史、视图等, 而是根据自己方案的特殊性, 将全部可能的泄露信息用泄露函数进行形式化表达, 这样做使得不同方案的安全性表述以及安全性证明更加清晰统一, 对可搜索加密安全性的研究工作具有重大意义。不过, Chase 所做的安全模型改进仍然是建立于静态 SSE 方案之上。

4.2.3 引入动态性后最终的适应性安全模型

2012 年, Kamara 等人^[17]第一次将动态性正式引入 SSE 方案。相应地, Kamara 将之前 Chase 等人在静态环境下定义的适应性安全模型进行了扩展, 引入了新的泄露函数 L_{Update} 用于表达用户在更新文件的过程中所产生的信息泄露, 例如更新时间和更新文件的大小等。我们将 Kamara 最终得到的适应性安

全模型称之为 L -适应性安全模型,这也是目前公认的 SSE 需要满足的安全模型,我们给出它的完整定义如下^[17,26]。

一个泄露函数: $L = (L_{Setup}, L_{Search}, L_{Update})$ 形式化地包含了对称可搜索加密方案中建立、搜索、更新过程中产生的泄露信息,为定义 L -适应性安全首先需要定义两个模拟游戏,现实游戏 $Game_{R,A}$ 和理想游戏 $Game_{S,I,A}$ 如下:

在 $Game_{R,A}$ 中,通过给安全参数 λ , 给予敌手 A 通过 Setup 协议建立的数据结构 EDS , 而 $Game_{S,I,A}$ 中,给予敌手的 $EDS' = S(L_{Setup}(1^\lambda))$ 是由模拟器 S 通过泄露函数模拟的 EDS' 。敌手在两个游戏中进行适应性的搜索询问或者更新操作即可以根据上次的结果进行调整,在 $Game_{R,A}$ 中,敌手得到真实的搜索结果,而在 $Game_{S,I,A}$ 中则由模拟器模拟的数据作为结果即 $S(L_{Search}(EDS, q_i))$ 和 $S(L_{Update}(upd, w))$ 。最后敌手 A 返回一个比特 $0(Game_{R,A})$ 或 $1(Game_{S,I,A})$ 表示区分,具体细节可见图 4。

$Game_{R,A}$	$Game_{S,I,A}$
1: $(\sigma, (ED, I)) \leftarrow Setup(1^\lambda, \perp)$	1: $(\sigma, (ED, I)) \leftarrow S(L_{Setup}(1^\lambda))$
2: for $k = 1$ to q do	2: for $k = 1$ to q do
3: $Q_k = (type_k, para_k)$	3: $Q_k = (type_k, para_k)$
4: if $type_k = \text{update}$ then	4: if $type_k = \text{update}$ then
5: $R_k \leftarrow Update((\sigma, para_k), (ED, I))$	5: $R_k \leftarrow S(L_{Update}(para_k))$
6: else	6: else
7: $R_k \leftarrow Search((\sigma, para_k), (ED, I))$	7: $R_k \leftarrow S(L_{Search}(para_k))$
8: end if	8: end if
9: $A \leftarrow (Q_k, R_k)$	9: $A \leftarrow (Q_k, R_k)$
10: end for	10: end for
11: $b \leftarrow A$	11: $b \leftarrow A$
12: return b	12: return b

图 4 安全游戏

Figure 4 Security games

在理想游戏中敌手得到的结果都是由模拟器根据泄露信息所模拟出的结果,如果敌手 A 不能区分出这两个游戏,即通过泄露函数模拟的数据库与真实的数据库在操作中使得敌手无法做出区分。则说明方案除了泄露函数中包含的信息外,并无更多的信息泄露。因此,在以上条件下,我们只需要考虑泄露函数本身的信息的安全。根据文献[5]中的总结,以前大多数方案的泄露函数中包含 4.2.1 节中定义的访问模式和搜索模式等信息。然而在最终的适应性安全模型中, Curtmola 等人已经证明它们在被动敌手的假设下是足够安全的。另外,如数据库大小,关键词总数,时间信息等常见的泄露函数包含信息,通常被认为是可以被敌手拿到而不影响方案的安全。然而随着研究的发展,被动模型下某些证明安全的泄露信息成为了恶意敌手的攻击手段,具体见第 5 节。

定义 4.5.(L -适应性安全). 一个对称可搜索加密

系统满足 L -适应性安全(L -adaptive security)是指对所有可能的具有多项式时间算法的敌手 A , 对于以上定义的游戏模型存在一个模拟器 S 满足下列等式,其中 $negl(\lambda)$ 为可忽略函数:

$$|\Pr(Game_{R,A}(\lambda) = 1) - \Pr(Game_{S,I,A}(\lambda) = 1)| \leq negl(\lambda)$$

总而言之,研究者们以 Curtmola 的静态环境下的 SS 自适应安全模型作为雏形,在逐渐完善下,得到了 L -适应性安全模型。不同的方案中泄露函数有所区别,但形式上的统一为安全性方面的研究提供了很大的便利。直观上, L -适应性安全的定义可以参考图 5, 这一安全保证了 SSE 方案在被动敌手的安全假设下具有足够的安全性,因此也是所有 SSE 方案应当满足的最基本的安全模型。

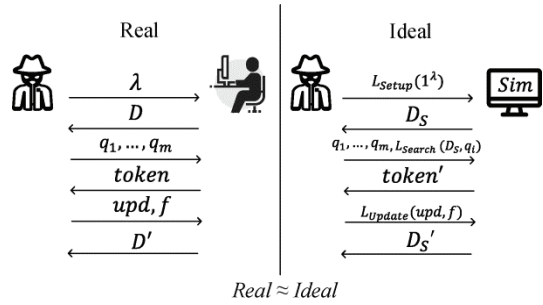


图 5 L -适应性安全

Figure 5 L -adaptive security

然而,随着近年来的研究发现,如第 5 节描述的,对称可搜索加密的安全性仍需进一步完善。在 L -适应性安全中被证明安全的泄露信息在主动敌手模型下不再安全。因此仅仅满足 L -适应性安全的 SSE 方案不能抵挡恶意敌手发起的一些新型的主动攻击手段,研究者们为此提出了新的安全性定义和解决方法。

5 对称可搜索加密安全性研究的新进展

如第 3, 4 节所述, L -适应性安全模型能够完全抵抗半诚实的被动敌手模型。但恶意敌手模型的引入导致对称可搜索加密的安全性出现了两类新的挑战,一类是 5.1 节提到的如 IKK 攻击等利用数据挖掘思想与公共数据库进行比对的攻击手段的提出,这里我们统称为推理攻击(Inference Attack),另一类是 5.2 节提到的文件注入攻击的威胁。

我们强调:两类攻击的目的都在于恢复用户搜索的关键字内容。而由于文件与关键字紧密联系,如果一份文件的多个关键字被泄露无异于文件内容本身失去隐私。如何解决这两种攻击揭示的安全问题,成为了目前对称可搜索加密安全性研究的重点方向。

5.3 节和 5.4 节将说明面对这两个新挑战, 目前研究者们所提出的新的安全定义以及解决方法。

5.1 推理攻击

由于研究的前沿性, 这类攻击的命名并未统一, 例如文献[23]称这类攻击为泄露滥用攻击(Leakage-abuse attacks), 在文献[27-28]中将他们提出的攻击手段称之为重建攻击(reconstruction attack), 但本文我们将列举出的文献[22-23, 28-33]中的攻击统一称为推理攻击(Inference Attack), 而我们将说明这些攻击实际上运用的是同一类攻击思想。首先我们用 Islam 提出的具体的 IKK 攻击方案^[22]为例来说明这类推理攻击的攻击思想。

在 2012 年, Islam 等人^[22]首次发现了已有的对称可搜索加密系统存在这种安全漏洞: 用户进行搜索时的行为模式和搜索习惯会为服务器带来额外可收集的信息。例如某用户多次检索关键字后, 敌手通过监视存储空间和交互信息, 可以将返回的加密文件集合记录并对每个搜索结果进行频率统计, 并与公开数据库或已泄露文档比较。

Islam 提出的具体的 IKK 攻击方案如下:

考虑一个单关键字搜索的可搜索方案, 在假设下, 敌手 Mallory 掌握用户与服务器间的交互信息。用户向服务器提交一系列的搜索序列 $Q = \langle Q_1, \dots, Q_l \rangle$, 对应的, 敌手观察到每个搜索用户提交的陷门 $Trapdoor_{Q_i}$ 和返回结果序列 $R = \langle R_{Q_1}, \dots, R_{Q_l} \rangle$ 。Mallory 的目标在于确认陷门所对应的具体关键字, 即关键字序列 $K = \langle K_{a_1}, \dots, K_{a_l} \rangle$, 分别对应 $Trapdoor_{Q_i} = K_i$ 。IKK 攻击中定义了一个单词并发矩阵 M : 其中每个 (i, j) 矩阵元素第 i 个元素和第 j 个元素出现在任意文件 $d \in D$ 中的概率, 即 $M_{i,j} = Pr[(K_i \in d) \wedge (K_j \in d)]$, 其中 d 为均匀分布。敌手通过对公开数据库或者从用户之前已泄露的数据文件进行分析得到具体的矩阵 M , 最后可以规约为解决以下优化问题来破解关键字:

$$\underset{\langle a_1, \dots, a_l \rangle}{\operatorname{argmin}} \sum_{Q_i, Q_j \in Q} \left(\frac{R_{Q_i} \cdot R_{Q_j}^T}{n} - (K_{a_i} \cdot M \cdot K_{a_j}^T) \right)^2$$

参考 Islam^[22]中的实验结果, 结果表明 IKK 攻击具有相当高的成功率, 在关键字空间小于 2500 个时, 攻击还原关键字的成功率为 60%~80%, 且用户提出的搜索请求越频繁, 攻击的成功率就越高。此外, Islam 还在文章中表示这一攻击具有自适应性, 即敌手可以根据已猜测出的关键字对矩阵 M 进行调整,

达到更好的攻击效果。

和 IKK 攻击类似, 推理攻击一般是敌手通过数据挖掘, 收集到用户进行搜索和访问时的统计数据, 获得例如关键字搜索返回文件集合的大小等信息。之后, 敌手可以通过构建一些特征比较函数来比较用户数据与公开数据库间的特征相似之处, 不断推理出用户的关键字信息。在第 4 节描述的适应性安全模型无法抵挡这种攻击原因在于, 适应性安全模型是半诚实的被动敌手模型, 其中我们往往只考虑某一次单独的搜索或更新中的信息泄露是否安全, 但敌手不会主动收集用户的多次搜索或更新的历史信息进行利用。

因此这一类型攻击的特点在于通过用户多次, 大量的进行搜索或更新操作, 以此提供返回文件信息和陷门信息给敌手进行分析, 恢复出用户搜索的关键字内容。而且, 一旦某些关键字内容被破解或用户已经提前泄露了部分关键字内容, 将形成“多米诺骨牌”效应, 攻击的效果将大大增强。

在这之后, 2014 年, Liu 等人^[29]给出了一种新的推理攻击手段, 与 IKK 攻击的区别在于没有采用单词并发矩阵, 而是通过公开的单词频率统计工具得到单个关键字在公众数据库中的频率特征, 然后同样与用户进行多次交互, 收集并统计每个陷门对应的结果信息进行对比。

进一步地, 2015 年, Cash^[23]提出了计数攻击(Count Attack), 结合了之前 Islam 和 Liu 提出攻击的优点, 并改进了由于关键字空间增大会导致攻击效果显著下降的缺点如图 6 所示, 即在关键字空间较大时, Cash 的计数攻击仍有较高的成功率和效率。

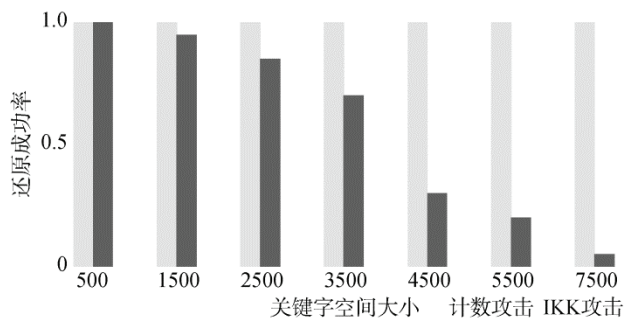


图 6 计数攻击与 IKK 攻击对比

Figure 6 Counting attack vs. IKK attack

近年来, 这种针对用户访问模式进行数据挖掘思想的推理攻击仍然不断被提出^[12, 30-34], 例如 Naveed 等人^[30]和 Grubbs 等人^[32]将推理攻击应用到满足性质保持(property-preserving)的加密数据库上给出了具体的攻击方案, Pouliot 等人^[31]则针对采用

了 EDESE(文件后附带加密关键字列表)设计思想的可搜索加密协议进行攻击, 结合图论中的 WGM(Weighted Graph Matching)问题来设计用户数据与公开数据库的比对方法。Kellaris 等人^[27]和 Lacharité 等人^[28]则将他们提出的攻击手段称之为重建攻击(reconstruction attack), 它们都是对支持范围查询(range queries)的加密数据库进行攻击, 攻击基本原理是标记范围查询中最小(或最大)的结果, 利用查询记录返回的频率结合已知的其他记录来进行标记和比对, 实际上我们可以看出这与之前所描述的推理攻击的基本思想是一致的, 都是按照数据收集-公开比对的模式来恢复关键字内容。

5.2 文件注入攻击

2016 年, Zhang^[24] 针对动态对称可搜索加密方案提出了文件注入攻击。这一攻击只针对具有动态性的 SSE 方案, 并且具有代价小, 攻击效果强的特点。

与 5.1 节中的推理攻击方式不同, 这一攻击方法是: 敌手首先向动态的对称可搜索加密系统注入一些含有特定关键字(敌手设计)的文件, 系统在更新这些文件后, 敌手根据用户原来提交过的用户陷门私自进行关键字查询, 通过查询结果还原陷门对应的关键字。

下面用一个简单的例子来说明文件注入攻击。假设关键字总共有 8 个分别编号 1~8, 敌手构建 3 份含有特定关键字的注入文件, 每份文件含有 4 个不同的关键字, 如图 7 所示, 第一份文件含有 4, 5, 6, 7 号关键字, 第二份文件含有 2, 3, 6, 7 号关键字, 第三份文件含有 1, 3, 5, 7 号关键字。用户以前查询过 2 号关键字并对应提交过陷门 T , 敌手为了还原 T 所对应的关键字, 首先向系统注入这 3 份特定文件并等待系统更新后, 根据陷门 T 私下对系统进行搜索。敌手根据三份注入的文件是否响应搜索, 如果响应结果为 0, 1, 0, 根据二分法的思想, 就可以确定陷门 T 对应了 2 号关键字, 从而破解了搜索的隐私安全。



图 7 文件注入攻击示例

Figure 7 Example of file-injection attack

文件注入攻击简单高效, 敌手可以任意设定可

能的关键字空间, 如果敌手事先了解用户的使用背景就可以限定到一个更小的关键字空间使得攻击更为成功。并且, 对于 K 个关键字, 只需要构造 $\log K$ 份注入文件即可, 代价很小。由于这一攻击对不满足前向安全的 DSSE 方案具有致命的威胁, 可以 100%地还原用户的关键字信息, 因此前向安全的可搜索加密方案得到了迅速的关注和深入的研究, 在 5.3.2 中进一步介绍。

5.3 新的安全模型和对抗手段

如 5.1 节与 5.2 节中所描述的, 对称可搜索加密目前主要面临两个新的安全性挑战。而在这些攻击提出以来, 可搜索加密的研究者们也积极地提出了新的安全模型和对抗手段。

5.3.1 针对推理攻击

根据推理攻击的原理, 近年来, SSE 的研究者们一般给出的解决办法是隐藏返回结果的大小(Result Pattern Hiding)。由于推理攻击一般是利用用户查询操作后, 返回文件时, 由于服务器对交互信息和存储空间的监视, 可以获得返回结果文件的个数与大小等信息, 以此和公共数据进行比对。因此, 隐藏返回结果的大小是直接且有效的方法。

从细节上, 目前隐藏返回结果的大小有填充(padding)和分区(partition)两种通用方法。填充是指对每个关键字返回的结果用虚假数据进行填充使得每个关键字的返回结果大小一致; 分区是指将同一个关键字在服务器不知情的情况下拆分成多个区域进行搜索, 以此隐藏单个关键字结果的真实大小。显然填充需要额外的存储和交互开销, 分区搜索需要额外的密钥管理和搜索时间的延长, 这都会对方案的效率产生很大影响。

还有一些其他的方法进行结果隐藏, 例如 2018 年, Lai 等人^[34] 就结合 HVE 方法 (Hidden Vector Encryption)给出了隐藏返回结果大小的支持多关键字查询的 SSE 方案。2019 年, Kamara 等人^[35]给出了一个通用转换器, 将填充和分区两种方法结合使用, 能够将一般的 SSE 方案进行转换达到隐藏返回结果大小的目的。虽然这些方案都能直接抵抗推理攻击, 但如何构建更为高效的方案或给出新的方法也是目前研究的重要方向之一。

5.3.2 针对文件注入攻击

2014 年, 针对动态对称可搜索加密, Stefanov^[26]首次提出, 在用户进行添加文件时应当满足前向安全, 在删除文件时应该满足后向安全。简单来讲, 前向安全是指添加的新文件不应该被以前的陷门查询到, 后向安全是指删除后的文件不能被服务器访问

到。虽然有部分文章^[37-38]对后向安全进行了一些研究,但目前为止还没有发现不满足后向安全所带来的安全漏洞,因此目前的研究重心主要在前向安全。

在 2016 年提出的文件注入攻击大大强调了前向安全的重要性。实际上满足前向安全就能够抵抗文件注入攻击,这是因为如 5.2 中所述,服务器要成功恢复关键字内容,必须用旧的陷门得到新注入文件的响应,而前向安全不允许旧的陷门查询到新的更新文件。因此近年来,前向安全作为新的安全模型得到了广泛的研究。

Stefanov 构造了一个满足前向安全的 DSSE 方案,却并没有给出这个安全概念的正式定义或讨论其重要性。在之后, Bost 等人^[20]才第一次正式形式化地定义了前向安全,并给出了一个前向安全的方案。

定义 5.1(前向安全) op 为添加或删除操作, ind 代表此更新文件标识符, u 代表此更新文件的大小, L' 为某泄露函数。一个可搜索加密方案是前向安全的当且仅当更新泄露函数 L_{upt} 可以被写为 L' :

$$L_{upt}(op, ind) = L'(op, (ind, u))$$

这一定义说明更新时的泄露信息只能严格包含操作类型、文件标识符、文件大小,而没有其他任何信息,因此 L_{upt} 可以被写成如上形式。实际上,前向安全并不与 L -适应性安全冲突,应当理解为在 L -适应性安全的基础上同时满足前向安全。之后, Bost 等人^[36]又提出了同时满足前向安全和后向安全的通用方案, 2017 年, Kim 等人^[21]设计了一种新的索引结构双重字典(dual dictionary), 以此给出了支持高效更新的前向安全方案。2018 年, Ghareh 等人^[37]的前向安全方案是利用伪随机函数生成器将更新文件时的密钥进行更新。同年, Mishra 等人^[38]结合 ORAM 技术思想给出了一个提供不经意访问的安全索引结构 Obliv 来实现前向安全。

5.4 利用 ORAM 思想进行设计

近年来,面对新的安全挑战,一些方案^[26,39-40]结合 ORAM 思想进行设计,以牺牲一定的效率来保证安全性。

Oblivious RAM (简称 ORAM) 技术是 1996 年^[41]提出用于云存储服务并进行数据保密的一种技术,是另一独立的研究方向,至今仍有很高的研究热度。ORAM 因为其完全隐藏用户的访问和存储信息而可以提供云存储中最高级别的安全。然而 ORAM 的劣势相当明显,即计算复杂度和带宽开销过大,效率极低。因此在第 5 节中描述新的安全问题提出以前,将 ORAM 用于 SE 的设计是冗余且低效的。文献^[26,39-40]中的方案没有直接利用 ORAM 结构作为

方案的主体,而是间接采用其“混淆”存储位置和访问信息的思想进行构造,实现了效率和安全的权衡,具有一定的可行性,但还需要进一步的效率提升。

6 结论

综上所述,自 2000 年对称可搜索加密被提出以来,到 2012 年 Kamara 等人完成了动态对称可搜索加密的 L -适应性安全模型的建立,对称可搜索加密的安全性研究已经逐渐趋向成熟。然而,2012 年以来,对称可搜索加密的安全性有了两个新的挑战:一是如何抵抗类似 IKK 攻击的推理攻击;二是对文件注入攻击和前向安全的进一步研究。针对前者,隐藏返回结果的大小是较为直接的方法,但这一方法会导致效率受到很大影响,还需要进一步的讨论和研究。在前向安全方面,虽然研究已经具备一定的成果,各种方法实现前向安全的方案层出不穷,但是效率和开销问题仍然需要进一步的取舍。除此之外,不满足后向安全具有的安全威胁还未被发现,值得未来探讨。

一般地,在构造可搜索加密方案时,研究者们需要考虑三个方面:安全性、表达能力和效率开销。安全性毋庸置疑是可搜索加密提出的前提,也是最重要的问题。在安全性的研究以外,效率和表达能力的研究也备受关注。表达能力主要考虑的是在进行检索询问时是否支持更多的功能如模糊搜索,连词搜索等。效率开销问题则是在前两个问题基础上将可搜索加密技术应用到现实设备中的一大阻碍。其中,如研究用户除了搜索单个关键字外还支持布尔查询的方案^[42-43],支持模糊查询^[44]等,通过结合 ABE 或 IBE 实现细粒度的访问控制的^[45-46]可搜索加密方案,对可搜索加密效率与安全间权衡问题的研究^[47-48]等。实际上,表达能力和效率的研究与安全性也是息息相关的。

从可搜索加密提出的最根本目的出发,为了保障用户使用云存储服务时的隐私安全,安全性问题依旧是可搜索加密技术未来研究最重要的一环。除此之外,可搜索加密技术仍然依附于现有的传统加密技术进行设计,而传统加密技术目前面临的一大挑战是量子计算机的诞生,虽然目前抗量子密码算法设计工作正如火如荼的进行,但还没有具体应用到可搜索加密技术中。

最后,总结一下可搜索加密技术如今面临的问题以及未来的发展方向:

1. 如何更有效地抵抗类似 IKK 攻击等推理攻击;
2. 对前向安全和后向安全进一步的探讨;

3.在安全性实现的同时如何权衡效率和表达能力以应用于实际;

4.抗量子的可搜索加密方案的讨论与设计。

参考文献

- [1] Shen Z R, Xue W, Shu J W. Survey on the Research and Development of Searchable Encryption Schemes[J]. *Journal of Software*, 2014, 25(4): 880-895.
(沈志荣, 薛巍, 舒继武. 可搜索加密机制研究与进展[J]. *软件学报*, 2014, 25(4): 880-895.)
- [2] Li J W, Jia C F, Liu Z L, et al. Survey on the Searchable Encryption[J]. *Journal of Software*, 2015, 26(1): 109-128.
(李经纬, 贾春福, 刘哲理, 等. 可搜索加密技术研究综述[J]. *软件学报*, 2015, 26(1): 109-128.)
- [3] Bösch C, Hartel P, Jonker W, et al. A Survey of Provably Secure Searchable Encryption[J]. *ACM Computing Surveys*, 2015, 47(2): 1-51.
- [4] Fuller B, Varia M, Yerukhimovich A, et al. SoK: Cryptographically Protected Database Search[EB/OL]. 2017: arXiv:1703.02014[cs.CR]. <https://arxiv.org/abs/1703.02014>
- [5] Poh G S, Chin J J, Yau W C, et al. Searchable Symmetric Encryption[J]. *ACM Computing Surveys*, 2017, 50(3): 1-37.
- [6] Dong X L, Zhou J, Cao Z F. Research Advances on Secure Searchable Encryption[J]. *Journal of Computer Research and Development*, 2017, 54(10): 2107-2120.
(董晓蕾, 周俊, 曹珍富. 可搜索加密研究进展[J]. *计算机研究与发展*, 2017, 54(10): 2107-2120.)
- [7] Li Y, Ma C G. Overview of Searchable Encryption Research[J]. *Chinese Journal of Network and Information Security*, 2018, 4(7): 13-21.
(李颖, 马春光. 可搜索加密研究进展综述[J]. *网络与信息安全学报*, 2018, 4(7): 13-21.)
- [8] Boneh D, di Crescenzo G, Ostrovsky R, et al. Public Key Encryption with Keyword Search[M]. *Advances in Cryptology - EUROCRYPT 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004: 506-522.
- [9] Byun J W, Rhee H S, Park H A, et al. Off-Line Keyword Guessing Attacks on Recent Keyword Search Schemes over Encrypted Data[M]. *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2006: 75-83.
- [10] Tang Q, Chen L Q. Public-Key Encryption with Registered Keyword Search[M]. *Public Key Infrastructures, Services and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010: 163-178.
- [11] Meng R, Zhou Y W, Ning J T, et al. An Efficient Key-Policy Attribute-Based Searchable Encryption in Prime-Order Groups[M]. *Provable Security*. Cham: Springer International Publishing, 2017: 39-56.
- [12] Liu X Y, Li H G, Zhang F G. Dynamic Searchable Encryption Scheme on Cloud Storage with Multi-level Access[J]. *Journal of Cryptologic Research*, 2019, 6(1): 61-72.
(刘翔宇, 李会格, 张方国. 一种多访问级别的动态可搜索云存储方案[J]. *密码学报*, 2019, 6(1): 61-72.)
- [13] Song D X, Wagner D, Perrig A. Practical Techniques for Searches on Encrypted Data[C]. *SP '00: The 2000 IEEE Symposium on Security and Privacy*. 2000: 44-55
- [14] Chang Y C, Mitzenmacher M. Privacy Preserving Keyword Searches on Remote Encrypted Data[J]. *Applied Cryptography and Network Security*, 2005: 442-45. DOI:10.1007/11496137_30.
- [15] Naveed M, Prabhakaran M, Gunter C A. Dynamic Searchable Encryption via Blind Storage[C]. *SP '14: The 2014 IEEE Symposium on Security and Privacy*. 2014: 639-654.
- [16] Curtmola R, Garay J, Kamara S, et al. Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions[J]. *Journal of Computer Security*, 2011, 19(5): 895-934.
- [17] Kamara S, Papamanthou C, Roeder T. Dynamic Searchable Symmetric Encryption[C]. *The 2012 ACM conference on Computer and communications security - CCS '12*, 2012: 965-976.
- [18] Goh E J. Secure Indexes[J]. *IACR Cryptol. EPrint Arch.*, 2003: 1-19.
- [19] Stefanov E, Papamanthou C, Shi E. Practical Dynamic Searchable Encryption with Small Leakage[C]. *NDSS*. 2014, 71: 72-75.
- [20] Bost R. Σοφος: Forward Secure Searchable Encryption[C]. *The 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016: 1143-1154.
- [21] Kim K S, Kim M, Lee D, et al. Forward Secure Dynamic Searchable Symmetric Encryption with Efficient Updates[C]. *The 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017: 1449-1463.
- [22] Islam M S, Kuzu M, Kantarcioglu M. Access Pattern Disclosure on Searchable Encryption: Ramification, Attack and Mitigation[J]. *Ndss*, 2012, 20: 1-15.
- [23] Cash D, Grubbs P, Perry J, et al. Leakage-Abuse Attacks Against Searchable Encryption[C]. *The 22nd ACM SIGSAC Conference on Computer and Communications Security*, 2015: 668-679.
- [24] Zhang Y, Katz J, Papamanthou C. All Your Queries Are Belong to Us: The Power of File-Injection Attacks on Searchable Encryption[C]. *USENIX Security Symposium*. 2016: 707-720.
- [25] Chase M, Kamara S. Structured Encryption and Controlled Disclosure[M]. *Advances in Cryptology - ASIACRYPT 2010*. Springer Berlin Heidelberg, 2010: 577-594.
- [26] Stefanov E, Papamanthou C, Shi E. Practical Dynamic Searchable Encryption with Small Leakage[C]. *NDSS*. 2014, 71: 72-75.
- [27] Kellaris G, Kollios G, Nissim K, et al. Generic Attacks on Secure

- Outsourced Databases[C]. *The 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016: 1329-1340.
- [28] Lacharité M S, Minaud B, Paterson K G. Improved reconstruction attacks on encrypted data using range query leakage[C]. *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018: 297-314.
- [29] Liu C, Zhu L H, Wang M Z, et al. Search Pattern Leakage in Searchable Encryption: Attacks and New Construction[J]. *Information Sciences*, 2014, 265: 176-188.
- [30] Naveed M, Kamara S, Wright C V. Inference Attacks on Property-Preserving Encrypted Databases[C]. *The 22nd ACM SIGSAC Conference on Computer and Communications Security - CCS '15*, 2015: 644-655.
- [31] Pouliot D, Wright C V. The Shadow Nemesis: Inference Attacks on Efficiently Deployable, Efficiently Searchable Encryption[C]. *The 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016: 1341-1352.
- [32] Grubbs P, Sekniqi K, Bindschaedler V, et al. Leakage-abuse attacks against order-revealing encryption[C]. *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2017: 655-672.
- [33] Kellaris G, Kollios G, Nissim K, et al. Generic Attacks on Secure Outsourced Databases[C]. *The 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016: 1329-1340.
- [34] Lai S Q, Patrnanabis S, Sakzad A, et al. Result Pattern Hiding Searchable Encryption for Conjunctive Queries[C]. *The 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018: 745-762.
- [35] Kamara S, Moataz T. Computationally Volume-Hiding Structured Encryption[M]. *Advances in Cryptology – EUROCRYPT 2019*. Cham: Springer International Publishing, 2019: 183-213.
- [36] Bost R, Minaud B, Ohrimenko O. Forward and Backward Private Searchable Encryption from Constrained Cryptographic Primitives[C]. *The 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017: 1465-1482.
- [37] Ghareh Chamani J, Papadopoulos D, Papamanthou C, et al. New Constructions for Forward and Backward Private Symmetric Searchable Encryption[C]. *The 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018: 1038-1055.
- [38] Sun S F, Yuan X L, Liu J K, et al. Practical Backward-Secure Searchable Encryption from Symmetric Puncturable Encryption[C]. *The 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018: 763-780.
- [39] Mishra P, Poddar R, Chen J, et al. Oblix: An efficient oblivious search index[C]. *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018: 279-296.
- [40] Miers I, Mohassel P. IO-DSSE: Scaling Dynamic Searchable Encryption to Millions of Indexes by Improving Locality[C]. *The 2017 Network and Distributed System Security Symposium*, 2017.
- [41] Goldreich O, Ostrovsky R. Software Protection and Simulation on Oblivious RAMs[J]. *Journal of the ACM*, 1996, 43(3): 431-473.
- [42] Cash D, Jarecki S, Jutla C, et al. Highly-Scalable Searchable Symmetric Encryption with Support for Boolean Queries[M]. *Advances in Cryptology – CRYPTO 2013*. Springer Berlin Heidelberg, 2013: 353-373.
- [43] Kamara S, Moataz T. Boolean Searchable Symmetric Encryption with Worst-Case Sub-linear Complexity[M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2017: 94-124.
- [44] Li J, Wang Q, Wang C, et al. Fuzzy keyword search over encrypted data in cloud computing[C]. *2010 Proceedings IEEE INFOCOM*. IEEE, 2010: 1-5.
- [45] Wang C J, Li W T, Li Y, et al. A Ciphertext-Policy Attribute-Based Encryption Scheme Supporting Keyword Search Function[M]. *Cyberspace Safety and Security*. Cham: Springer International Publishing, 2013: 377-386.
- [46] Xu P, Wu Q, Wang W, et al. Generating Searchable Public-Key Ciphertexts with Hidden Structures for Fast Keyword Search[EB/OL]. 2015: arXiv:1512.06581[cs.CR]. <https://arxiv.org/abs/1512.06581>
- [47] Asharov G, Segev G, Shahaf I. Tight Tradeoffs in Searchable Symmetric Encryption[M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018: 407-436.
- [48] Demertzis I, Papadopoulos D, Papamanthou C. Searchable Encryption with Optimal Locality: Achieving Sublogarithmic Read Efficiency[M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018: 371-406.



刘文心 于 2018 年在北京航空航天大学数学专业获得理学学士学位。现在北京航空航天大学数学专业攻读硕士学位。研究领域为可搜索加密、安全多方计算。研究兴趣包括：云存储。
Email: liuwenxin@buaa.edu.cn



高莹 于 2003 年在武汉大学数学与统计专业获得博士学位。现任北京航空航天大学网络空间安全学院副教授，博士生导师。研究领域为安全多方计算、生物特征加密、区块链、可搜索加密。研究兴趣包括：量子密码。Email: gaoying@buaa.edu.cn