

可搜索加密协议的访问模式保护技术研究综述

龚云平^{1,2,3}, 李雄^{1*}, 文博³, 武春岭², 童世华²

(1. 电子科技大学计算机科学与工程学院(网络空间安全学院), 四川成都 611731; 2. 重庆电子科技职业大学人工智能与大数据学院(网络空间安全学院), 重庆 401331; 3. 熊猫电子集团有限公司研究院, 江苏南京 210013)

摘要: 可搜索加密协议使用户能够对云端密文执行关键词检索而不泄露明文, 已成为云存储数据安全的关键技术。然而, 多数方案在查询执行中会泄露访问模式, 即每次查询返回的匹配文档集合及其对应关系。近年来研究表明, 攻击者利用访问模式泄露可发动查询恢复、数据库重建乃至文件注入攻击。本文对可搜索加密协议的访问模式保护技术进行系统综述。首先构建了以访问模式为核心的泄露谱系模型, 刻画了包括访问模式、搜索模式、体量模式等核心泄露类型的层次关联与联合利用关系; 其次提出四维技术分类框架, 从 ORAM/OMAP、混淆与差分隐私、分布式多服务器、硬件可信执行环境四条路线对代表性方案进行分类剖析与深度对比; 进而绘制攻防对抗演进图谱, 揭示攻击技术的升级规律; 在此基础上建立涵盖安全性、效率、功能、可部署性的多维评估指标体系与综合评价矩阵, 并补充工业落地现状与开源实现调研; 最后展望后量子安全、新型计算架构、高级搜索功能融合等方向的挑战与机遇。

关键词: 可搜索加密协议; 访问模式泄露谱系模型; 不经意随机访问机; 前向隐私; 后向隐私

Survey on Access Pattern Protection Techniques for Searchable Encryption Protocols

GONG Yunping^{1,2,3}, LI Xiong^{1*}, WEN Bo³, WU Chunling², TONG Shihua²

1. School of Computer Science and Engineering (School of Cyberspace Security), University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China
2. School of Artificial Intelligence and Big Data (School of Cyberspace Security), Chongqing Polytechnic University of Electronic Technology, Chongqing 401331, China
3. Research Institute, Panda Electronics Group Co., Ltd., Nanjing, Jiangsu 210013, China

Abstract: Searchable encryption protocols enable users to perform keyword searches on encrypted data in the cloud without revealing plaintext, making them a key technology for securing cloud-stored data. However, most existing schemes leak access patterns during query execution—specifically, the sets of matching documents returned by each query and their corresponding relationships. Recent studies have shown that attackers can exploit such leaked access patterns to launch query reconstruction, database reconstruction, and even file injection attacks. This paper provides a systematic review of access pattern protection techniques in searchable encryption. First, we establish a leakage lineage model centered on access patterns, characterizing the hierarchical relationships and combined exploitation among core leakage types—including access patterns, search patterns, and volume patterns. Second, we propose a four-dimensional technical classification framework, categorizing and conducting in-depth comparisons of representative schemes along four search directions: ORAM/OMAP, obfuscation and differential privacy, distributed multi-server architectures, and

Key words: searchable encryption protocol, access pattern protection, oblivious random access machine (ORAM), for-

收稿日期: 2026-08-02; 修回日期: XXXX-XX-XX

通信作者: 李雄, lixiong@uestc.edu.cn

基金项目: 目: 重庆市教委科学技术研究项目(KJQN202503108, KJQN202403116); 国家自然科学基金资助项目(No. U25A20429, No. 62332018)

Foundation Items: The Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN202503108 and No. KJQN202403116); The National Natural Science Foundation of China (No. U25A20429, No. 62332018).

ward privacy, backward privacy

hardware-based trusted execution environments. Third, we map out an evolution graph of attacker-defender dynamics, revealing the progression of attack techniques. Based on this foundation, we develop a multidimensional evaluation framework encompassing security, efficiency, functionality, and deployability, along with a comprehensive assessment matrix, supplemented by surveys on industrial deployment status and open-source implementations. Finally, we discuss future challenges and opportunities in areas such as post-quantum security, novel computing architectures, and integration with advanced search functionalities.

0 引言

随着云计算与大数据技术的深度渗透,越来越多的个人用户和企业机构选择将海量数据外包至云平台,以获取弹性的存储管理和高效的数据访问服务。然而,数据外包使数据的所有权与管理权发生分离,云端服务器通常并非完全可信,因此,数据必须以密文形式存储以保障机密性^[1]。传统加密技术能够有效保护数据内容的机密性,却导致数据丧失了基本的可计算性,从而在密文上执行检索操作变得极为困难。为解决该“密文可用性”悖论,可搜索加密 (Searchable Encryption, SE) 技术应运而生,它允许用户在不解密数据的前提下,对云端密文执行关键词检索而不泄露明文信息,达到了数据安全性和可用性的有效平衡。自 2000 年 Song 等人^[2]提出首个可搜索加密协议以来,该领域在二十余年间经历了从概念构建到功能丰富,再到隐私深化的演进历程。尽管现有的可搜索加密协议^[3-5]在支持动态更新、布尔查询等高级功能上取得了长足进步,但其在设计上仍以查询效率和功能丰富性为首要目标,而对查询过程中暴露的访问模式并没有构建系统性保护。

该技术的早期研究^[2,6]主要聚焦于可行性验证与基础构造,通过伪随机函数、布隆过滤器 (Bloom Filter) 等密码学原语在单关键词场景下实现了亚线性搜索,之后逐步扩展到支持动态更新^[3-4]、联合关键词搜索^[5-7]、排序搜索^[8]、模糊搜索^[9]等更为复杂的功能。然而,这些方案在设计之初对信息

泄漏的认识相对有限,默认服务器仅能获取密文索引和查询陷门。随着研究深入,研究者逐渐意识到,为了维持搜索效率,绝大多数方案不可避免地向外泄露“访问模式”,即查询所返回的匹配文档集合及其对应关系^[10-11]。该局面随着 Islam 等人^[12]于 2012 年首次通过实验证实访问模式泄漏可被用于高精度查询恢复而发生了根本性改变,文中宣称至少 80% 的密文关键词可被正确恢复。此后,泄漏滥用攻击不断涌现,从早期的统计推断被动攻击^[12-13]发展到基于文件注入的主动攻击^[14-16],再到如今的融合搜索模式、体量模式等多维泄漏的联合攻击^[17-19]。这些工作深刻揭示了访问模式泄漏的严重危害,加速推动可搜索加密协议研究领域进入以隐私保护为核心的新阶段。

在早期的可搜索加密协议中,服务器通过观察陷门访问的索引位置和返回的文档列表,可以轻易地关联两次相同关键词的查询、推断关键词的文档频率乃至重建整个数据库索引^[20]。这种单点功能导向的设计模式在处理半可信甚至恶意服务器时暴露出显著的局限性:方案满足传统的语义安全定义,却未能提供对访问模式该核心泄漏源的任何理论保障。为了解决该问题,访问模式保护技术作为可搜索加密的一个独立而关键的研究分支被明确提出并快速发展,其目标是在维持密文检索可用性的前提下,通过密码学原语、混淆机制或架构创新,实现对查询执行过程中暴露的文档集合及其关联信息的不可区分。这些技术能够模拟“查询无痕”的理想安全状态,从而处理远超传统方案防护能力的泄漏滥用攻击,助力实现云存储数据“可用不可见”的最终目标。该挑战与密码学自身的发展逻辑不谋而合,即为了对抗不断增强的攻击能力,防护技术必须从单点加密向系统级隐私保护演进。

在理想的安全定义下,可搜索加密协议应只泄露访问的“输出”即匹配文档集合,而不应泄露任何关于“过程”的额外信息。然而,为了实现亚线性甚至常数级的搜索效率,绝大多数实用方案在倒排索引查询、陷门匹配、结果返回等环节,会向服务器暴露程度不等的信息,这些信息被系统化为“泄漏谱”。其中,访问模式是泄漏谱中最核心、危险性最高的组成部分,它揭示了每次搜索查询所触及的加密文档标识集合及其对应关系^[21-22]。Islam

等人^[12]于2012年首次通过实验证实,利用访问模式泄漏,攻击者能够以较高概率正确恢复用户的密文搜索查询内容。该发现犹如打开了潘多拉魔盒,此后十余年间,围绕访问模式的攻防研究持续演化,形成了一场激烈的技术博弈。目前,对可搜索加密的一般性综述已较为丰富^[11,23-25],然而对可搜索加密协议的访问模式保护技术领域的系统性梳理待进一步完善。

为了明确该领域的独特研究脉络,我们强调具备访问模式保护的^{可搜索加密协议}与传统^{可搜索加密协议}之间的三个核心区别。首先是安全目标的升级:传统方案的设计目标是保证明文内容和陷门的语义安全,其泄漏谱是方案执行后被动产生的副产品;与之相对,访问模式保护方案将“隐藏查询访问了哪些文档”作为显式的主动设计目标,从源头控制信息暴露。这种模式将安全分析者的角色从“泄漏的测量者”转变为“泄漏的消除者”。其次是敌手模型的增强:早期的SSE安全分析通常假设一个“诚实但好奇”的服务器,而访问模式保护技术为应对更强大的敌手,包括能够执行文件注入的主动攻击者^[14]和利用多泄漏源联合推断的统计攻击者^[19],提供了专门的防御手段。最后是研究重心从效率优先向安全-效率权衡的工程化迁移:经典SSE研究的核心在于降低搜索复杂度(如达到 $O(1)$ 或 $O(\log N)$),而访问模式保护技术的研究显著地转向了在信息论安全(如ORAM^[26])、统计上安全(如差分隐私^[27])和可调节泄漏(如SEAL^[28])之间寻找可部署的平衡点,这涉及通信开销、在线交互轮数、存储膨胀和功能兼容性等多维工程约束的精细化折衷^[29]。

攻击与防御之间呈现出典型的“螺旋递进”演化规律。一方面,攻击技术日趋精密:从早期基于频率分析的被动攻击^[12-13],发展到利用文件注入的主动攻击^[14-16],再到融合搜索模式、体量模式、结果长度模式等多维泄漏的联合攻击^[17-19],甚至出现了利用大语言模型生成合成知识来增强攻击能力的新型范式^[30-31]。另一方面,防御技术也从简单的填充混淆^[32]演进为复杂的ORAM (Oblivious Random Access Machine) 集成^[33-35]、差分隐私混淆^[36-37]、分布式数据分片^[38-40]、硬件可信执行环境保护^[41-42]等多条技术路径。如今,访问模式保护技术正在对云存储数据安全实践产生深远影响。以

OSSE^[37]和PANCAKE^[36]为代表的混淆方案,已能在常数倍通信开销下提供形式化的统计保护;而基于分布式架构^[38]和可信执行环境^[41]的方案则为特定场景提供了高安全性的替代路径。但是,目前仍存在一些重要问题没有得到妥善解决。首先,完备保护方案与真实云环境的集成困难重重,其 $O(\log^2 N)$ 的带宽膨胀比在大规模高并发服务中难以承受。其次,现有统计保护方案主要针对被动攻击设计,面对注入攻击或LLM增强攻击^[30]的防御能力尚不充分,迫使安全设计者不断升级混淆策略。

近年来,访问模式保护已迅速跃升为可搜索加密领域最为活跃的研究分支之一。目前,已有的综述^[24-25,43-44]多侧重于可搜索加密的通用安全定义或功能分类,这在一定程度上模糊了访问模式保护该专项领域多样技术路线背后共通的设计逻辑。软件学报此前发表的可搜索加密综述^[11,43]对SE的基础机制和分类体系进行了详尽的介绍;吴鹏飞等人的综述^[45]系统梳理了ORAM原语的构造、优化与性能分析,但该综述聚焦ORAM本身,没有面向可搜索加密场景,没有针对SE特有的访问模式、搜索模式、体量模式等泄漏开展系统梳理,也未讨论SE场景下ORAM与倒排索引集成带来的组合谬误等特有问題;都未对可搜索加密协议的访问模式保护技术领域的系统性梳理待进一步完善。此外,随着大语言模型、可信执行环境(Trusted Execution Environment, TEE)、后量子密码等新技术的涌现,访问模式的攻防格局正在经历深刻重塑,有必要对该领域最新的研究动态进行系统梳理并展望未来方向。

为此,该工作旨在专门研究可搜索加密协议的访问模式保护技术,以期致力于产学研的工作者提供参考。首先,系统回顾了可搜索加密协议的演化历程,构建了以访问模式为核心的泄漏谱系模型,定量刻画了包括访问模式、搜索模式、体量模式等核心泄漏类型的层次关联与联合利用关系;其次,提出四维技术分类框架,从ORAM/OMAP、混淆与差分隐私、分布式多服务器、硬件可信执行环境四条技术路线出发,对可搜索加密协议的访问模式保护代表性方案进行了分类剖析和深度对比;同时,绘制了攻击技术与防御策略的对抗演进图谱,揭示了从被动统计推断到LLM增强主动注入的攻击升级规律;在此基础上,建立了涵盖安全

性、效率、功能、可部署性的多维评估指标体系和综合评价矩阵；最后，前瞻了该领域在后量子安全、新型计算架构、高级搜索功能融合和面向大语言模型时代攻防演进等研究方向的挑战和机遇。

本文的组织结构如下：首先，建立可搜索加密协议的泄漏谱系模型，阐明访问模式在其中的核心地位与危害；其次，深入探讨可搜索加密协议中访问模式保护的四条关键技术路线；再次，系统梳理攻击与防御的演进博弈规律，并建立多维评估指标体系；最后，分析该领域当前面临的挑战并展望未来方向，对全文进行总结。

1 可搜索加密协议的基础与演化历程

1.1 基本概念与系统模型

可搜索加密是一种密码学原语，允许数据所有者将加密数据外包至半可信的云端服务器，同时保留对密文数据执行关键词搜索的能力^[2]。如图1所示，一个典型的对称可搜索加密 (Symmetric searchable encryption, SSE) 系统涉及三个参与方：数据所有者、服务器和用户。

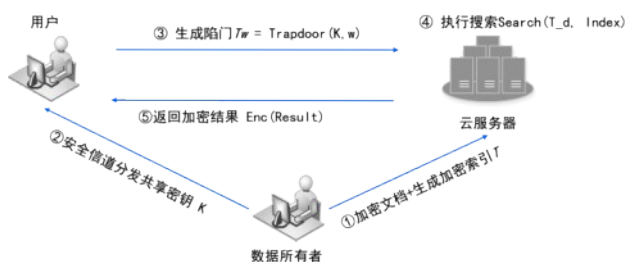


图1 对称可搜索加密系统模型

在上述系统中，数据所有者先对明文文档集合 $D = \{d_1, d_2, \dots, d_n\}$ 进行加密处理，同时基于关键词空间 $W = \{w_1, w_2, \dots, w_m\}$ 构建安全的加密索引 T 。加密文档与加密索引一同上传至服务器。当用户需要搜索包含关键词 w 的文档时，它利用与所有者共享的密钥 K 生成搜索陷门 T_w 并发送给服务器。服务器在不知晓 w 明文的前提下，利用 T_w 在 T 上执行搜索操作，最终返回匹配的加密文档标识符集合。整个过程中，服务器作为“诚实但好奇”的敌手，严格遵循协议执行，但试图从交互过程中推断出尽可能多的敏感信息。

从密码学构造的维度，可搜索加密协议可划分为两大范式：基于对称密码的 SSE^[2,10,46] 和基于公钥密码的 PEKS (Public Key Encryption with Key-

word Search)^[47-48]。前者通常利用伪随机函数、对称加密和抗碰撞哈希函数等高效对称原语构建，搜索效率较高，适用于单写者/多读者场景；后者则基于公钥基础设施构建，支持多写者/单读者场景，但计算开销通常更大。本文重点聚焦于实用化程度更高的 SSE 方案，但多数访问模式保护思想可被推广至 PEKS 领域。

在索引结构层面，为支持亚线性复杂度的搜索，绝大多数 SSE 方案采用倒排索引作为基础数据结构^[10]。倒排索引为每个关键词 $w \in W$ 维护一个匹配文档列表 $DB(w) = \{id(d) : w \in d\}$ ，搜索时只需定位到目标关键词对应的列表即可。然而，这种高效的结构也恰恰成为泄漏的根源，是服务器通过观察搜索陷门命中的索引位置以及返回的文档列表，可以轻易地推断出查询之间的关联关系和结果集统计特征。

1.2 可搜索加密协议的演化四阶段

可搜索加密协议的发展可以清晰地划分为四个代际，每个代际在功能支持、安全定义和效率平衡上呈现出鲜明的特征。如表1所示，从多个维度总结了这四个演化阶段的核心特点。

第一阶段：概念形成与基础构建。 Song 等人^[2]首次展示了在对称加密数据上进行关键词搜索的可行性，方案采用双层加密结构，搜索过程需逐词比对，计算复杂度为 $O(n)$ 。Curtmola 等人^[10]的工作是该阶段的集大成者，他们给出了 SSE 的第一个严格安全性定义 (非自适应 / 自适应语义安全)，并提出了基于倒排索引的高效构造方案 SSE-1 和 SSE-2，奠定了 SSE 方案设计的基石。

第二阶段：功能扩展与安全保障。 研究者开始关注更接近实际应用场景的搜索功能。联合关键词搜索^[5-6]允许用户使用多个关键词的布尔表达式进行查询；排序搜索^[8]根据相关性对结果进行排序；模糊搜索^[9]容忍拼写错误和不精确匹配。同时，支持动态更新的 DSSE (Dynamic Symmetric Searchable Encryption) 方案成为研究热点。Kamara 等人^[3]提出了首个支持添加和删除操作的 DSSE 方案。Cash 等人^[5]提出的 OXT 协议是里程碑式的工作，它在大规模数据库上高效支持布尔查询，并成为后续大量学术研究和工业部署的基础，如 Microsoft 的 CryptDB 部分思想。然而，该阶段虽然功能大为增强，但方案设计者普遍未将“访问模式泄漏”视

为严重安全威胁。

第三阶段：攻击揭示与隐私觉醒。Islam 等^[12]在 2012 年提出的泄漏滥用攻击，在 2015 年之后随着多篇高影响力攻击工作的问世而引发广泛关注。Cash 等人^[6]和 Zhang 等人^[14]分别展示了即使符合当时安全性定义的 SSE 方案，但在访问模式泄漏面前依然脆弱。该严峻挑战推动了隐私保护意识的全面觉醒。在防御方面，Bost 等人^[49,76]形式化了前向隐私、后向隐私的定义，为 DSSE 设定了更高安全基准；同时，ORAM 开始被系统性地引入 SE 领域，从而不经意多映射(Oblivious multi-Map, OMAP)^[33]被提出。

第四阶段：攻防博弈与多维演进。攻击与防御进入持续的深度博弈阶段。攻击方面，LLM 辅助攻击^[30-31]、体积注入攻击^[50-51]、多泄漏源联合攻击^[18-19]不断涌现，攻击所需背景知识假设被持续降低。防御方面，混淆与差分隐私保护^[36-37,52]、分布式多服务器架构^[38-40]、TEE 硬件方案^[41-42]等多条技术路线并行发展，呈现百花齐放态势。同时，研究者^[29]对现实部署中的延迟容忍度、渐进部署能力、存量系统兼容性等工程约束也日益重视。

1.3 前向隐私与后向隐私

在动态对称可搜索加密协议中，前向隐私和后向隐私是衡量动态更新安全性的两个核心概念。理解它们与访问模式保护之间的区别与联系，对于系统化掌握 SE 的安全全景至关重要。

定义 1.1 (前向隐私)：若一个 DSSE 方案满足前向隐私，则对于任意插入操作，服务器无法将该插入与之前任何已执行的搜索查询相关联^[49, 53]。更准确地说，新添加的文档及其关键词索引不会泄露与先前搜索陷门的任何匹配关系。

定义 1.2 (后向隐私)：后向隐私限制删除操作的信息泄露。根据泄露程度的递增，后向隐私泄露分为如下的三个等级^[49]：

Type-I：仅泄露删除操作发生的时间和当前数据库大小的变化；

Type-II：在 Type-I 基础上，额外泄露被删除的文档标识符属于哪一个关键词列表；

Type-III：在 Type-II 基础上，额外泄露每次删除操作对应的具体（关键词，文档 ID）对。

Bost^[53]通过基于陷门置换的方案实现了前向安全；Sun 等人^[7]和 Ghareh Chamani 等人^[54]分别基于可穿刺加密 (Puncturable Encryption) 提出了支持不同程度后向隐私的实用方案。

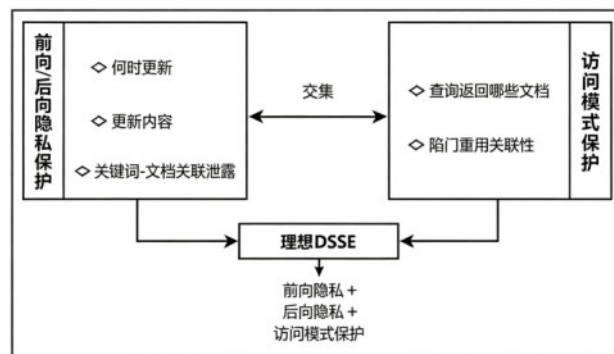


图2 访问模式保护与前向/后向隐私的关系示意图

访问模式保护与前向/后向隐私的异同，如图2所示，这两组安全概念聚焦于不同的泄漏维度，但存在重要的交集。前向隐私和后向隐私主要解决更新信息，即何时添加/删除了什么的泄漏问题，而访问模式保护则聚焦于查询结果层面的信息暴露，即每次查询返回了哪些文档。两者是正交但互补的安全维度。一个方案可能同时完美满足前向和后向

表 1

可搜索加密协议演化四阶段对比

维度	第一阶段	第二阶段	第三阶段	第四阶段
核心目标	可行性验证	功能丰富化	攻击揭示与隐私觉醒	攻防博弈与多维演进
典型搜索功能	单关键词线性扫描	联合查询、排序与模糊搜索	前向/后向安全更新	支持复杂表达式的隐私保护查询
动态性	仅静态	支持动态更新	关注更新隐私	更新与查询隐私协同保护
安全定义	基本语义安全	基本泄漏刻画	前向隐私、后向隐私定义	统计保护、可调节泄漏
泄漏认知	几乎未被关注	初步描述泄漏谱	泄漏滥用攻击频发	多维泄漏联合建模
代表性方案	[2,10]	[3,5]	[49,53-54]	[36-38,41]
技术特征	顺序扫描、布隆过滤器	倒排索引、OXT 协议	ORAM 引入、陷门置换	差分隐私、TEE、分布式

隐私, 却仍然在访问模式上存在严重泄漏, 从而遭受查询恢复攻击^[17]。反之, 某些访问模式保护方案, 如基础的 ORAM-SSE, 可能未显式提供前向/后向隐私保证。因此, 在工程实践中, 通常需要将两类技术进行有机组合, 以实现全方位的信息泄漏控制。

2 泄漏谱系分析与建模

为了准确评估可搜索加密协议的访问模式保护技术的效能与必要性, 本节对可搜索加密中面临的泄漏信息进行系统建模, 厘清各类泄漏的定义、层次关系及其潜在危害, 为后续的技术分类与评估奠定理论根基。

2.1 泄漏类型的形式化定义

在可搜索加密协议的语境下, 泄漏指方案在正常执行过程中, 不可避免地向服务器暴露的、非直接输出信息^[10-11]。根据泄漏信息的性质和粒度, 可将其归纳为以下几种主要类型。如表 2 所示, 给出了这些泄漏类型的定义、符号表示及数学描述。

定义 2.1 (搜索模式) : 搜索模式 L_{sp} 揭示任意两次搜索查询是否对应于同一个关键词。形式化地, 对查询序列 $Q = \{q_1, q_2, \dots, q_n\}$, 搜索模式定义了一个 $n \times n$ 等价关系矩阵 M , 其中 $M[i, j] = 1$ 当且仅当 q_i 和 q_j 搜索的是相同关键词, 否则为 0。搜索模式泄露了用户查询行为的重复性与关联性^[10, 18]。

定义 2.2 (访问模式) : 访问模式 L_{ap} 揭示每次搜索查询返回的匹配文档标识符集合。对于查询 q 搜索关键词 w , 访问模式^[12, 55, 77] 定义为 $L_{ap}(q) = \{id(d) : d \in D \wedge w \in d\}$, 其中 $id(d)$ 是文档 d 的唯一标识符。访问模式直接暴露了“哪些文档包含我所搜索的关键词”的核心关联信息。

定义 2.3 (体量模式) : 体量模式 L_{vp} 揭示每次查询返回的匹配文档数量。定义为: $L_{vp}(q) = |L_{ap}(q)|$, 体量模式^[19] 可以被视为访问模式的一个粗粒度“投影”, 它丢失了具体的文档身份信息, 但保留了关键的基数信息。

定义 2.4 (结果长度模式) : 结果长度模式 L_{rp} 揭示每次查询返回的加密文档总字节长度^[51]。它与体量模式相关但不完全等同, 因为不同文档的加密后长度可能存在差异, 从而泄露更丰富的信息。

定义 2.5 (更新模式) : 在动态方案中, 更新模式 L_{up} 揭示添加或删除操作发生的时间戳和操作

类型 (添加/删除), 但可能在理想情况下不泄露受影响的文档与关键词之间的具体关联^[34]。

定义 2.6 (缓存模式) : 在更强大的威胁模型中, 如存在恶意系统管理员或利用硬件侧信道, 服务器可能观察到数据访问过程中的内存地址序列或磁盘块访问序列, 即缓存模式 L_{cp} 。该泄漏是比逻辑上的访问模式更底层的物理信息^[56]。

定义 2.7 (ORAM 的不经意性) : 令客户端发起的真实访问序列为 $\vec{a} = (a_1, a_2, a_3, \dots, a_M)$, 其中每个 $a_i = (op_i, addr_i, data_i)$ 表示对标记为 $addr_i$ 的数据块内容 $data_i$ 进行一次读或写操作 op_i , 读操作表示对数据内容取回来后未更新, 而写操作表示对数据内容取回来后进行了更新。对于除了用户自己外, ORAM 不仅可保证被访问的数据块之间不可区分, 而且可确保每个数据块的操作也是不可区分。因此, 形式化如下: ORAM 协议将 $\vec{a}$ 转换为物理访问序列 $Access(\vec{a})$, 使得对于任意两个等长的访问序列 $\vec{a}$ 和 $\vec{b}$, 物理访问序列的统计分布不可区分: $Access(\vec{a}) \stackrel{c}{\approx} Access(\vec{b})$, 其中 $\stackrel{c}{\approx}$ 表示计算上 / 统计上 / 信息论上不可区分性。

2.2 泄漏层次与蕴含关系

上述多种泄漏类型并非彼此独立, 而是呈现出明确的层次结构和信息蕴含关系。图 3 以有向图的形式刻画了主要泄漏类型之间的信息流与涵盖关系。

图中的箭头方向表示“信息蕴含”关系。例如, 访问模式 L_{ap} 直接蕴含体量模式 L_{vp} , 因为可以从文档 ID 集合的基数推导出匹配数量; 同时, 在大多数静态 SSE 方案中, 若两次查询的访问模式完全相同, 则可大概率推断它们对应同一关键词 (除非关键词的文档分布完美冲突), 因此访问模式也在很大程度上蕴含搜索模式的信息。反之, 体量模式或搜索模式单独则不能重建完整的访问模式。最底层的缓存模式包含了最丰富的信息, 它通常是 ORAM 等技术需要对抗的物理泄漏源。

2.3 可搜索加密协议中访问模式泄漏的潜在危害

在泄漏谱系中, 访问模式之所以被公认为危害最大的核心泄漏, 是因为它可能被用于多种高破坏力的攻击, 直接威胁用户查询隐私和密文数据库重建。

(1) **查询恢复攻击**。这是最常见和研究最充分

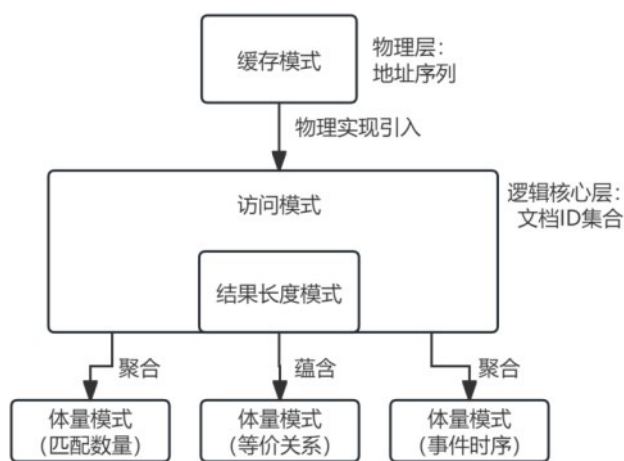


表2 主要泄漏类型的形式化定义、符号表示及数学描述

的一类攻击。攻击者拥有部分背景知识，如某些用户文档的明文版本、关键词的全局统计分布等，通过将观测到的访问模式矩阵与背景知识进行统计匹配，从而恢复用户查询的关键词^[12-13]。Islam 等人^[12]在 Enron 邮件数据集上实验显示，当攻击者知晓约 20% 的明文文档时，单关键词查询的恢复率可超过 80%。

(2) 数据库重建攻击。在更持久的威胁模型下，攻击者可以通过长期积累的访问模式信息，逐步推演整个加密数据库的文档-关键词索引结构^[20,57]。这实际上等同于“恢复”了数据库的核心内容关联，其危害远超单次查询的暴露。

(3) 文件注入攻击。作为一种主动攻击模型，文件注入攻击允许攻击者向数据所有者的数据集植入精心构造的“标记文件”，随后通过观察后续查询的访问模式中是否包含这些注入文件，以极高置信度确定查询的关键词^[14-16]。Zhang 等人^[14]提出的 ZKP 攻击仅需 $\log_2|W|$ 个注入文件即可破解后续所有查询，是目前最具威胁性的攻击范式之一。

(4) 统计推断与复合攻击。访问模式与体量模

式、搜索模式等泄漏的组合使用，使得攻击者能够进行更精确的统计推断。特别是 Oya 和 Kerschbaum^[17]的突破性工作表明，即使访问模式被信息论安全，如通过 ORAM，单独的搜索模式泄漏仍可被用于高精度查询恢复。该发现极大地扩展了泄漏攻击的视野，也强调了防御必须兼顾多类泄漏源。

2.4 访问模式保护安全等级术语界定

(1) 信息论安全：也叫不经意安全，以 ORAM/OMAP 为代表，敌手无论计算能力多强，均无法从协议执行过程获取访问模式信息，可抵抗被动攻击；

(2) 统计安全：以差分隐私、混淆填充类方案 (OSSE、PANCAKE) 为代表，基于统计不可区分，在隐私预算约束下保护访问模式，抵御被动统计攻击，但难以对抗主动注入攻击；

(3) 计算安全：敌手为概率多项式时间敌手，在计算复杂度假设下实现隐私保护。

3 访问模式保护技术分类与剖析

根据可搜索加密协议中访问模式保护的核心技术原理和所依赖的安全假设，本文将现有的方案系统归纳为四条主要技术路线，如表 3 所示。本节将对每条路线进行深入的机理剖析、代表性方案详述和性能分析。其中的代表性方案主要选取 2013-2025 年间访问模式保护方向代表性工作；其中 2013 年 Path ORAM 属于该领域奠基性经典工作，虽时间超过 10 年，但作为后续大量 SE 访问模式保护方案的基础原语，纳入综述对象；筛选维度包含：发表于顶级会议或期刊、提出全新机制、具备原型实现、对后续研究有显著影响力的论文。本节中的泄漏谱系模型完成“问题刻画”，系统拆解 SE 协议中搜索模式、访问模式、体量模式等多类

图3 泄漏类型之间的层次关系与信息蕴含图

泄漏类型	符号表示	定义	信息粒度	数据来源
搜索模式	L_{sp}	两次查询是否对应相同关键词	等价关系矩阵	陷门观察
访问模式	L_{ap}	每次查询返回的匹配文档集合	文档 ID 集合	搜索结果
体量模式	L_{vp}	每次查询返回的匹配文档数量	标量	结果集大小
结果长度模式	L_{rp}	返回加密文档的总长度 (字节数)	标量	通信量观察
更新模式	L_{up}	更新操作的时间、类型 (增/删)	事件序列	更新交互
缓存模式	L_{cp}	服务器内存/磁盘访问序列	地址序列	硬件 / OS 观察

泄漏,明确不同泄漏的危害层级;四维技术分类框架完成“解决方案归类”,四类技术路线分别针对泄漏谱系中不同泄漏源、不同安全目标进行防护。ORAM/OMAP 目标消除逻辑层乃至底层缓存模式泄漏;混淆与差分隐私路线针对访问模式、体量模式做统计层面抑制;分布式多服务器通过多方拆分,切断单服务器获取完整泄漏视图;TEE 硬件路线在硬件隔离域阻断外部敌手观察访问相关泄漏。每一类技术路线可以防护泄漏谱系中一个或者多个泄漏类型,部分方案仅缓解部分泄漏。

3.1 路线 I:基于 ORAM / OMAP 的全隐藏路线

3.1.1 ORAM 核心思想与性能瓶颈

不经意随机访问机 (Oblivious Random Access Machine, ORAM) 由 Goldreich 和 Ostrovsky^[58]首次提出,旨在隐藏程序对远程存储器的访问模式。其核心思想是在每次访问数据时,对数据进行重新加密和随机重排,使得服务器观察到的物理访问序列与逻辑访问序列不可区分。以 Path ORAM^[26]为例,其是最具影响力的 ORAM 构造之一。它将数据块组织为一棵二叉树,Path ORAM 的带宽膨胀比为 $O(\log^2 N)$,其中 N 为用户外包到云端的总数据块数,在典型参数下约为明文查询的 $10^2 \sim 10^3$ 倍。TWO RAM^[59]首次将 ORAM 通信轮数优化至仅需两轮,其构造将客户端准备加密查询和服务器执行不经意访问并返回加密结果完全分离,与标准 SSE 的双轮交互模式完美契合。

3.1.2 OMAP 框架与 SE 集成挑战

将 ORAM 原语直接应用于 SSE 并非易事。简单地将 ORAM 用于 SSE 的倒排索引访问,不能直接提供对搜索模式和访问模式的全面保护,因为 SSE 的搜索操作涉及对索引中多个不连续位置的非

顺序访问,而标准 ORAM 被设计为单次随机访问原语。

为克服上述障碍,不经意多映射 (Oblivious Multi-Map, OMAP) 框架^[33-34]应运而生。OMAP 将 SSE 的倒排索引抽象为“关键词→文档列表”的多映射数据结构,并严格定义了三个层次的操作原语: Get (读取一个关键词对应的整个文档列表)、Put (追加一个文档 ID 到关键词列表)、Delete (从关键词列表中移除一个文档 ID)。OMAP 在每个原语内部调用 ORAM,实现对该映射的完全不经意访问,从而同时隐藏访问模式和搜索模式。

OMKX^[33]是 OMAP 的代表性实例化方案,它基于 Path ORAM 构建了支持多映射的不经意操作集,其搜索复杂度为 $O(a \cdot \log^2 N)$,其中 a 为匹配文档数。后续工作^[34,60]通过引入压缩列表、批量路径读取等优化技术,进一步将实际通信开销降低至可接受水平。PATCONFDB^[61]展示了 ORAM 在数据库索引保护上的另一个维度。该方案专门针对外包数据库的 B-tree 索引,实现了与具体 ORAM 构造解耦的“ORAM 无关”设计,使系统能够灵活适配不同的底层 ORAM 实现,在上千万条记录规模上额外增加的查询延迟控制在合理范围内。

3.2 路线 II:基于差分隐私的统计保护路线

鉴于 ORAM / OMAP 路线的高宽带开销和在线交互成本,部分研究者另辟蹊径,探索利用混淆机制和差分隐私 (Differential Privacy, DP) 提供形式化的统计保护,在安全性与效率之间寻求更优的平衡点^[27,36-37,62]。

3.2.1 差分隐私在访问模式保护中的应用

差分隐私由 Dwork^[63]提出,是隐私保护领域的黄金标准,它保证任意单个数据项的修改对统计输

表 3 访问模式保护技术四维分类框架

路线 I	路线 II	路线 III	路线 IV
ORAM / OMAP	混淆与差分隐私	分布式多服务器	硬件可信执行环境
Tree ORAM / Path ORAM	差分隐私、频率平滑、伪随机洗牌、可调节泄漏	原语: 密码共享、不经意传输、分布式 ORAM	原语: Intel SGX, ARM TrustZone, AMD SEV
优势: 信息论安全	优势: 常数级开销,可调安全参数	优势: 降低在线交互、专业化效率优化	优势: 通用计算、存量系统可改造
缺陷: 高宽带、高存储	缺陷: 统计安全非绝对安全	缺陷: 非共谋假设、部署模型苛刻	缺陷: EPC 内存限制、侧信道风险
代表方案: [33, 26, 59, 61]	代表方案: [37, 36, 28, 69]	代表方案: [38, 39, 67, 68]	代表方案: [41, 42, 70]

出的影响被严格控制在预设的隐私预算 (ϵ, δ) 之内。将其引入 SSE 访问模式保护,意味着通过对访问模式注入受控噪声,使得攻击者无论拥有何种背景知识,都无法以高置信度推断任何单一文档-关键词关联。

Chen 等人^[27]提出首个具备差分隐私访问模式保护框架。其方案在每次查询返回的结果集中,以一定概率 $p = f(\epsilon)$ 随机添加虚假文档或抑制真实文档,从而构建一个“加噪版本”的访问模式。该框架提供形式化的 (ϵ, δ) -DP 保证,然而其直接实现会引入大量噪音,严重降低搜索的精确度,且无法隐藏搜索模式。

3.2.2 OSSE 与 PANCAKE:两种代表性混淆范式

OSSE (Obfuscated Searchable Encryption)^[37]是对上述挑战的直接回应。OSSE 的核心创新在于为每次查询独立生成混淆后的访问模式,使得两次搜索相同关键词所表现出的服务器侧访问模式完全不同。由于搜索模式通过独立混淆被自然地隐藏,OSSE 对 IKK^[12]、计数攻击^[13]等经典被动攻击表现出极强的防御效果。其通信开销为 $O(a + \lambda * m)$,其中 a 为匹配文档数, m 为关键词空间大小, λ 为安全参数。

PANCAKE^[36]则从另一个角度切入,提出了“频率平滑”技术。其核心思想是将关键词的实际文档频率分布转化为均匀分布,使得任何两次查询在体量和访问模式层面都呈现出完全相同的统计特征。在与 ORAM 方案的对比实验中,PANCAKE 在同等安全假设下的吞吐量通常高出 1-2 个数量级。

SEAL^[28]提出了一种“可调节泄漏”的概念,允许系统设计者通过旋钮参数在安全性和查询延迟之间自由调节。方案通过对返回结果进行智能分组、填充和随机重排,在保留部分查询效率的同时有效降低了访问模式的信息泄漏量。

3.3 路线 III:基于分布式架构的多服务器路线

如果在威胁模型中引入多个非共谋的独立服务器,则可以利用分布式计算协议实现效率更高的访问模式保护^[39-40,64-65]。双服务器模型是最常见且实用的分布式假设。客户端将加密数据分片存储在两个互不共谋的服务器上。查询时,客户端与两个服务器分别进行一轮交互,每个服务器仅获取部分查询结果或加密的计算中间值。由于没有任何单一服务器能获取完整的访问模式视图,泄漏被有效遏制

在单个服务器的局部信息内。

OSM-DSSE^[38]是双服务器模型下的代表性高效方案。它基于洗牌算法和 Paillier 同态加密结合 1-out-of- n 不经意传输协议,在不使用传统 ORAM 的情况下实现了访问模式的隐藏。该方案的搜索效率较同类 ORAM 方案提升 3 至 4 倍,在处理高达 9×10^{10} 个关键词-文件对时的搜索时间仅为 2.19 秒。DOD-DSSE^[67]则基于分布式不经意数据结构,在双服务器模型下构造了专门的 DSSE 方案。通过将倒排索引的树形结构分布式存储,并设计双服务器协同的不经意遍历协议,方案避免了通用 ORAM 的高昂开销。端到端 SSE 方案^[68]提出了基于修改 Path-ORAM 的动态端到端 SSE 方案,通过双服务器协作实现访问模式隐藏。方案将加密文档拆分为主文档块和额外块,分别存储在两个服务器上,并对两个服务器的数据采用不同的混淆策略。这种设计使得任何一个服务器都极难通过响应身份识别攻击推断关键词-文档关联。该路线的最大局限在于“非共谋”假设的商业化云端部署可行性较低,通常需要依赖两个独立且互不信任的云服务商。

3.4 路线 IV:基于硬件可信执行环境的路线

可信执行环境 TEE 技术,尤以 Intel Software Guard Extensions (SGX)^[71]为代表,为访问模式保护提供一种全新硬件辅助解决方案^[41,72]。可信执行环境在 CPU 内部创建一个被称为“飞地”(Enclave)的加密隔离执行环境。“飞地”内的代码和数据即使对高权限的操作系统、Hypervisor 乃至 BMC 固件也是不可见的。在 SSE 部署中,搜索逻辑的核心部分如陷门解析、索引查找、结果解密等,在飞地内部执行,而不可信服务器仅负责提供加密数据的存储和 I/O 服务。服务器层面观察到的物理访问模式是由不可信 OS 管理的加密内存页面序列,“飞地”外观察者只能看到加密的、无顺序信息访问序列。

POSUP^[73]是实现上述理念的代表性方案。它利用 Intel SGX 构建高效的不经意数据结构,支持在超大规模数据集上进行不经意关键词搜索和更新。其关键在于通过将丰富但不可信的服务器计算资源与 SGX 提供的有限但安全的内存空间有机结合,实现高吞吐量。HardIDX^[41]采用一种更为务实的策略,即最小化可信计算基。方案仅将核心的倒排索引查找和结果过滤逻辑部署在“飞地”内,而

索引本身的存储和访问仍由不可信软件管理,减少了“飞地”内存的压力。SGX-MR-Prot^[42]则从易用性角度出发,提出了基于 MapReduce 框架的通用访问模式保护方案。它通过提供高层编程抽象,允许开发者几乎无需修改现有 MapReduce 应用程序代码即可获得访问模式保护能力。

上述 TEE 路线受制于“飞地”物理内存的有限容量,早期 SGX 仅约 128MB,处理超大规模数据时需要频繁页换入换出,可能导致严重性能下降;同时,多种侧信道攻击(如 Foreshadow^[57])已被证实可突破 SGX 的安全边界,泄露“飞地”内的访问模式。因此,TEE 路线在短期内的实用化仍面临挑战。

3.5 方案多维度对比与性能分析

为给出直观的技术路线全貌,如表 4 所示,从保护强度、安全假设、性能开销、功能支持等维度,对上述四条路线中的代表性方案进行了综合对

比。“保护强度”一列采用 2.4 节界定的统一术语:路线 I 方案统一标注为“信息论安全”,路线 II 方案统一标注为“统计安全”,路线 III 和路线 IV 方案标注为“信息论安全”。

路线 I (ORAM) 和路线 III (分布式) 能够达到统计上的全保护,但路线 I 的计算和通信开销最高,路线 III 依赖非共谋假设。路线 II (混淆/差分隐私) 在更低性能开销下提供有理论保障的统计保护,但无法达到绝对的访问模式全隐藏。路线 IV (TEE) 在全保护等级上与路线 I 相当,但受限于硬件假设和内存瓶颈。实际部署中,路线 II 因其在平衡点上的突出表现而最受工业界青睐。

从上述分析可以总结出四条技术路线在保护强度-效率图谱上的分布规律,如图 4 所示。

3.6 四条技术路线的横向比较与选型建议

综合上述分析,四条技术路线在安全性、效率、功能支持和可部署性四个维度上各有侧重,适

表 4

代表性方案对比

注: N 为数据块总数, a 为匹配文档数, κ 为安全参数, m 为关键词空间大小, τ 为频率平滑常数因子。

方案	技术路线	保护强度	安全假设	通信开销	在线交互	存储膨胀	动态支持	复杂查询
Path ORAM ^[26]	路线 I	信息论安全	单服务器诚实好奇	$O(\log^2 N)$	是(多轮)	$O(\log^2 N)$	有限	否
TWORAM ^[59]	路线 I	信息论安全	单服务器诚实好奇	$O(\kappa \cdot \log^2 N)$	2 轮	$O(\kappa \cdot \log^2 N)$	有限	否
OMKX ^[33]	路线 I	信息论安全	单服务器诚实好奇	$O(a \cdot \log^2 N)$	是(多轮)	$O(a \cdot \log^2 N)$	支持	否
OSM-DSSE ^[38]	路线 III	信息论安全	双服务器非共谋	$O(a)$	1 轮	$O(a)$	支持	否
DOD-DSSE ^[67]	路线 III	信息论安全	双服务器非共谋	$O(a \cdot \log N)$	多轮	$O(a \cdot \log N)$	支持	否
端到端 SSE ^[68]	路线 III	信息论安全	双服务器非共谋	$O(\log^3 N)$	多轮	$O(\log^3 N)$	支持	否
OSSE ^[37]	路线 II	统计安全	单服务器诚实好奇	$O(a + \lambda \cdot m)$	1 轮	$O(a + \lambda m)$	支持	否
PANCAKE ^[36]	路线 II	统计安全	单服务器诚实好奇	$O(\tau \cdot a)$	1-2 轮	$O(\tau \cdot a)$	支持	否
SEAL ^[28]	路线 II	统计安全	单服务器诚实好奇	可调节	1 轮	可调节	支持	否
Vaccine ^[69]	路线 II	计算安全	单服务器+注入攻击者	$O(a)$	1 轮	$O(a)$	有限	否
POSUP ^[73]	路线 IV	信息论安全	可信 SGX	$O(\log^2 N)$	是(多轮)	$O(\log^2 N)$	支持	否
HardIDX ^[41]	路线 IV	信息论安全	可信 SGX	$O(\log N)$	否	$O(\log N)$	有限	否

表 5

四条技术路线的选型建议

应用场景特征	推荐路线	理由
安全要求极高,可容忍延迟	路线 I (ORAM/OMAP)	信息论安全保障,可证明抵抗被动攻击
安全与效率需平衡	路线 II (混淆/差分隐私)	常数级开销,统计安全有理论保障
存在可信第三方或可部署多服务器	路线 III (分布式多服务器)	效率接近明文查询,依赖非共谋假设
已有 TEE 硬件基础设施	路线 IV (TEE)	对存量系统改造最小,性能可控
仅需抵御特定攻击(如注入攻击)	路线 II 中的定向方案	开销极低,针对性强



图4 保护技术路线在“保护强度-性能开销”图谱上分布

用于不同的应用场景。如表5所示，给出了基于具体应用需求的路线选型建议。

4 攻击防御的演进博弈

理解攻击技术的演进脉络，对于评估和设计访问模式防御机制至关重要。本节将系统梳理自2012年以来主要攻击方案的演进历程，并分析防御策略的相应发展，最后归纳博弈演化内在规律。

4.1 被动攻击技术

被动攻击指攻击者仅通过观察系统正常查询执行过程中泄露的信息来推断敏感数据，而不主动干预系统运行。IKK攻击^[12]是该领域的开山之作。Islam等人假设攻击者拥有部分用户的明文文档及其查询频率分布，将观测到的访问模式矩阵 $n_{queries} \times n_{docs}$ 与背景知识进行最优匹配，通过求解一个图匹配优化问题来恢复查询关键词。在Enron邮件数据集上，当背景知识覆盖率达到20%时，查询恢复率超过80%。计数攻击^[13]展示了仅利用体量模式泄露即可完成有效攻击。攻击者统计每次查询返回的文档数量，并将其与已知的关键词全局文档频率分布进行匹配。该方法简单高效，且不要求攻击者拥有任何具体文档的明文内容。利用部分文档集合作为背景知识，构造文档-查询二分图，并使用子图匹配技术来识别查询关键词^[17]。相对于IKK攻击，该方法在同背景知识水平下具有更高的查询恢复精度。

4.2 主动攻击技术

主动攻击允许攻击者在数据外包前或查询过程中向系统注入特制数据，从而获取更强的攻击能力。

4.2.1 已有代表性主动攻击技术

文件注入攻击是迄今为止最强大的一类主动攻击。攻击者通过共享文件夹、邮件附件等渠道向数据所有者的文档集合中植入精心构造的“标记文件”，随后在用户搜索时观察这些注入文件是否出

现在返回结果中。Zhang等人^[14]提出的ZKP攻击利用二进制编码的文件注入策略，仅需 $\log_2|W|$ 个注入文件，即可在后续的所有搜索中恢复查询关键词，攻击成功率趋近于100%。BVMA^[51]和Violin^[52]进一步深化了注入攻击的研究。它们发现，除了逻辑上的访问模式外，响应体量，即返回结果的加密数据长度，也是一种极易被利用的泄漏源。通过精心设计注入文件的长度编码方案，攻击者可以进一步减少所需的注入文件数量或提高攻击的隐蔽性。VAL攻击由Lambregts等人^[19]提出，它改进了一种名为LEAP的攻击方案^[74]的匹配技术，综合利用访问模式和体量模式进行攻击。其最大的特点是匹配算法不产生假阳性，且对已知文档的依赖度低，使得攻击结果更具确定性和破坏力。搜索模式独立攻击带来了认知上的重大转变。Oya和Kerschbaum^[17]证明了即使访问模式被完全隐藏，如通过部署ORAM，攻击者仅凭搜索模式泄露，即陷门的重复使用模式，仍然可以实现有效的查询恢复。他们证明了搜索模式中蕴含的关键词频率信息和共现信息，在合理的统计学假设下足以高精度地推断查询关键词。该发现意味着，访问模式保护是必要但不充分的，理想的SE方案必须同时隐藏搜索模式。

4.2.2 未来潜在方向

基于大语言模型（LLM）的合成数据攻击代表了攻击技术的前沿探索方向。LLM增强攻击^[30-31]代表了主动攻击的最新前沿。Chiu等人利用GPT-4等大语言模型生成合成数据来弥补真实背景知识的不足。通过精心设计的提示词，LLM可以生成统计特征高度逼近真实数据分布的合成文档集合，攻击者再使用传统统计匹配方法进行查询恢复。实验表明，该方法在仅拥有极少量真实数据泄露的情况下，也能达到与拥有大量真实背景知识的经典攻击相当的性能。需要指出的是，该方向尚处于早期探索阶段，目前仅有Chiu等人^[34]的初步工作，其攻击效果受限于合成数据的统计保真度，尚未形成成熟的技术体系。与文件注入攻击、搜索模式独立攻击等已形成完整方法论体系的攻击技术相比，LLM增强攻击在攻击成功率、适用条件和防御对抗性方面仍需进一步验证。然而，该方向展示了利用生成式AI增强泄漏攻击的潜力，值得后续研究持续关注。

4.3 防御策略的世代演进

面对攻击技术的持续升级,防御策略也经历了五个代际的深化发展。如图5所示,以时间线的形式展示了攻击与防御之间的关键博弈节点。

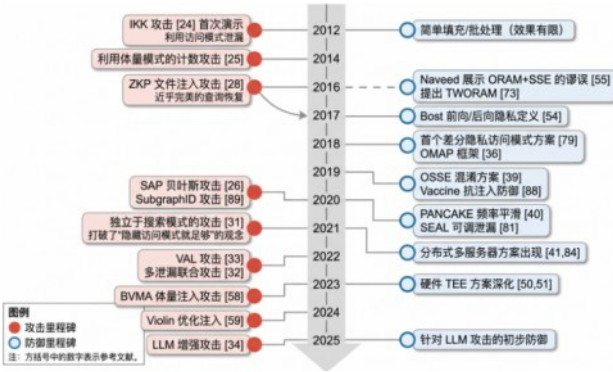


图5 可搜索加密访问模式攻防演进时间线

第一代防御(~2014): 简单填充与延迟返回。早期的方案^[32]尝试通过固定长度填充或批处理返回以混淆体量模式和访问模式。然而, Islam等人^[12]等人的攻击证明, 这种缺乏严格安全保障的启发式方法在面对统计攻击时效果非常有限, 无法提供实质性的保护。

第二代防御(2015~2019): ORAM集成与OMAP框架。随着Path ORAM^[26]、TWRORAM^[59]等高效ORAM构造的提出, 以及OMAP框架^[33]对ORAM与SSE组合问题的系统解决, 完全的访问模式隐藏成为可工程化的方案。虽然代价高昂, 但ORAM提供了可证明强安全性的保证。

第三代防御(2018~2021): 混淆与差分隐私引入。OSSE^[37]和PANCAKE^[36]等方案标志着防御路线的重要转向: 从追求“绝对安全”转向追求“理论保障的、开销可控的”统计安全。差分隐私的形式化保障、独立混淆的搜索模式隐藏、频率平滑的效率优化, 使得访问模式保护在逼近实用化的道路上迈出了一大步。

第四代防御(2019~2023): 定向防御与架构创新。Vaccine^[69]专门针对文件注入攻击设计, 通过“自我注入”策略混淆攻击者的观测, 不需要密码学强度级别的全保护即可有效抗击最强威胁。分布式多服务器方案^[38-39,66-67]通过引入非共谋服务器假设, 在效率与安全的天平上增加了新的平衡砝码。

第五代防御(2023至今): 硬件增强与智能防御。硬件TEE方案^[41-42,70]和面向LLM时代的智能

混淆策略成为新兴方向。利用机密计算硬件隔离执行搜索核心逻辑, 可以较低的计算开销实现内存中数据的明文处理安全; 同时, 利用生成式AI生成逼真的混淆查询来对抗LLM辅助的攻击, 开启了防御的智能化新维度。

4.4 博弈演化的内在规律

通过系统化地审视上述十余年的攻防对抗历程, 可以提炼出以下四条关键的演化规律:

(1) **螺旋递进律**: 攻击与防御呈现典型的“挑战-响应”螺旋式演进。一个世代的有效防御方案被提出后, 攻击者很快会找到其未覆盖的泄露源, 如从访问模式转到搜索模式; 或发掘更强的攻击假设, 如从被动到主动, 进而推动防御的再升级。

(2) **多维扩展律**: 攻击从早期的单一泄露源利用, 逐步演进为对访问模式、体量模式、搜索模式、响应长度等多个泄露维度的联合利用, 攻击能力与维度数量正相关。防御设计也必须由单点防护转向多维协同防护。

(3) **假设弱化律**: 攻击技术的演进方向是持续弱化对攻击者背景知识假设的依赖, 即从需要大量已知明文文档, 到仅需统计分布, 再到几乎不需真实数据而利用LLM合成数据。这大幅增加了防御的难度, 使得基于背景知识限制的早期安全论证不再可靠。

(4) **按需权衡律**: 面对性能开销的现实约束, 防御方案从追求“一刀切”的信息论安全, 转向更加务实的“按需保护”和“可调节安全”。针对特定攻击模型, 如Vaccine抗注入; 或特定应用场景, 如SEAL可调节泄露的定制化防御。

4.5 工业落地现状与挑战

近年来, 可搜索加密协议的访问模式保护技术正从学术研究走向工业实践。本节系统梳理该领域的开源实现与工业部署现状, 并分析落地面临的主要挑战。如表6所示, 为当前具有代表性的开源实现项目及其核心特征。

在医疗健康领域, 基于区块链与可搜索加密的电子病历共享方案已在多家医疗机构开展试点部署。这类方案利用云服务器存储加密电子病历, 通过可搜索加密技术实现跨机构的关键词检索与安全共享^[6-7]。在工业物联网领域, 面向云辅助工业物联网的高效可搜索属性基加密方案已在智能电网、智能制造等场景中得到验证, 支持多关键词搜索和

访问控制^[2]。在企业级应用^[4,9]中, CipherSweet 已被多个 PHP 项目采用, 用于实现数据库字段级加密与精确匹配搜索; Cosmian Findex 提供了 Rust 实现的多语言 API, 支持在零信任云环境对加密数据进行搜索。

尽管取得了上述进展, 访问模式保护技术的工业落地仍面临三个主要挑战。第一, 性能约束: 信息论安全方案 (路线 I) 的 $O(\log^2 N)$ 带宽膨胀比在大规模高并发服务中难以承受, 当前工业部署主要集中在统计安全方案 (路线 II) 上。第二, 安全假设现实性: 路线 III 依赖的“双服务器非共谋”假设在商业化云环境中难以验证, 路线 IV 依赖的 TEE 硬件存在侧信道风险且受厂商锁定限制。第三, 标准化缺失: 当前缺乏统一的 API 标准和互操作性规范, 不同方案之间的数据格式和协议接口不兼容, 增加了系统集成成本^[49]。

5 方案评估与多维指标体系

为了给方案选型和后续研究提供系统化的评价参照, 本节构建涵盖安全性、性能效率、功能支持、可部署性四个维度的综合评估指标体系, 并据此形成代表性方案的综合评价矩阵。

5.1 安全性指标

安全性指标评估方案对抗访问模式泄漏攻击的理论保障强度和实际抗攻击效能。

S_1 · 保护完备性: 衡量方案在理论层面达到的访问模式保护级别。分为三级: 信息论安全 (信息论意义下的不经意性)、统计保护 (提供 (ϵ, δ) -差分隐私或等效的形式化统计保障)、启发式保护 (无严格理论保证, 但在特定攻击模型下有效)。

S_2 · 抗被动攻击能力: 通过已知被动攻击方案 (IKK^[12]、LLM 增强^[30-31]等) 进行测试, 以查询恢复率和数据库重建率作为量化指标。恢复率越低, 说明保护能力越强。

S_3 · 抗主动攻击能力: 重点评估方案抵御文件注入攻击^[28,51-52]的能力。衡量指标包括攻击者达到一定恢复率所需的最小注入文件数、注入攻击的检测难度等。

S_4 · 搜索模式保护: 方案是否同时隐藏搜索模式。如前所述, 隐藏访问模式但暴露搜索模式仍可导致高精度攻击^[17], 因此该项是区分方案防护全面性的关键指标。

S_5 · 体量模式与长度模式保护: 方案在保护访问模式的同时, 是否顺带提供了对体量模式、结果长度模式的隐藏或混淆, 以抵御多维联合攻击。

5.2 性能效率指标

性能是决定方案能否走出实验室、迈向大规模部署的核心约束。

E_1 · 搜索计算复杂度: 包括服务器端和客户端的计算开销, 通常表示为总数据量 N 和匹配文档数 a 的函数的渐近形式。

E_2 · 通信开销与带宽膨胀比: 以明文查询所需通信量为基准, 计算保护方案的通信量放大倍数。ORAM 类方案通常为 $O(\log^2 N)$ 至 $O(\log^3 N)$ 倍, 混淆类方案多为常数至对数倍。

E_3 · 在线交互轮数: 方案执行一次搜索所需的客户端-服务器通信轮数。非交互式或单轮方案在网络高延迟环境中具有显著优势。

E_4 · 存储膨胀比: 保护方案在服务器端和客户端引入的额外存储开销。客户端暂存区 Stash 大小和服务器填充块数量是主要来源。

5.3 功能支持指标

F_1 · 动态操作支持: 是否支持高效的文档添加和删除, 以及添加/删除操作的隐私保护程度 (前向/后向隐私)。

F_2 · 高级查询支持: 是否支持布尔表达式查询、范围查询、子串查询、聚合查询等超越简单关键词匹配的高级搜索功能。

F_3 · 多用户与访问控制: 是否原生支持多用户场景下的细粒度访问控制和密钥分发。

5.4 可部署性指标

D_1 · 安全假设的现实性: 所依赖的安全假设 (如单服务器诚实好奇、双服务器非共谋、可信硬件等) 在实际商业云环境中的可实现性。

D_2 · 系统改造代价: 方案对现有云存储架构的修改范围和侵入性。客户端软件修改、服务器端索引重建等均会产生部署阻力。

D_3 · 依赖性: 方案对特定密码学库、硬件特性、或公钥基础设施的依赖程度。

5.5 综合评价矩阵

在 5.1 至 5.4 节中, 本文构建了涵盖安全性、性能效率、功能支持和可部署性四个维度的评估指标体系, 共包含 18 项具体指标 (S_1 - S_5 、 E_1 - E_4 、 F_1 - F_3 、 D_1 - D_3)。为避免评价过程的主观性和不可复现

性, 本节首先给出各指标的详细评分准则, 然后说明权重分配与综合得分计算方法, 最后基于统一的参数假设给出代表性方案的综合评价矩阵。

5.5.1 评分准则

如表 7 所示, 展示了全部 15 项指标从 1 星到 5 星的具体评分准则。每项指标的评分依据均为方案公开文献中可查证的设计特征、理论证明或实验数据, 评审者可根据统一准则复现评分。

5.5.2 权重分配与综合得分计算

各指标的重要性因应用场景而异。本文面向通用云存储场景, 采用层次分析法确定权重, 构造判断矩阵并通过一致性检验 ($CR < 0.1$)。如表 8 所示, 展示了维度权重。

表 8 维度权重		
维度	权重	依据
安全性 (S)	0.40	访问模式保护的核心目标
性能效率 (E)	0.30	决定方案能否实际部署
功能支持 (F)	0.20	影响方案适用范围
可部署性 (D)	0.10	工程落地约束

各维度内部指标采用等权处理, 则第 i 个方案的综合评分计算公式为:

$$\text{Score}_i = 0.40 \cdot \frac{1}{5} \sum_{j=1}^5 S_j^{(i)} + 0.30 \cdot \frac{1}{4} \sum_{j=1}^4 E_j^{(i)} + 0.20 \cdot \frac{1}{3} \sum_{j=1}^3 F_j^{(i)} + 0.10 \cdot \frac{1}{3} \sum_{j=1}^3 D_j^{(i)}$$

其中 $S_j^{(i)} \in \{1, 2, 3, 4, 5\}$ 为方案 i 在第 j 项安全性指标上星级评分, 依此类推。综合评分范围为 1-5 分。

5.5.3 综合评价矩阵

基于上述评分准则、权重和参数假设, 如表 7 所示, 展示了 10 种代表性方案在全部 15 项指标上的评分及综合得分。所有评分均由两位作者独立完成, 表 7 代表性访问模式保护方案综合评价矩阵分歧项经讨论后达成一致; 评分依据均来自方案原文中的理论证明或实验数据。

5.5.4 结果分析与讨论

如表 7 所示, 可以得出以下结论:

(1) 分布式方案综合表现最优: OSM-DSSE^[38] 以 4.42 分位居首位, 其在安全性和性能效率上的均衡表现突出, 主要失分项在于可部署性维度 ($D_1=3$, 依赖双服务器非共谋假设)。PANCAKE^[36] 以

4.22 分紧随其后, 其混淆路线在保持统计保护的同时实现了最低的通信开销和存储膨胀。

(2) TEE 方案性能突出但假设代价高: Har-dID^[41] 在 15 项指标中有 8 项满分, 综合分 4.18, 但 D_1 (假设现实性) 仅为 2 星, D_3 (依赖性) 仅为 2 星, 反映了其对特定硬件和厂商的强依赖。

(3) ORAM 路线安全性完备但效率评分偏低: Path ORAM^[26] 在 S_1-S_5 全部获得 5 星, 但 E_2 (通信开销) 和 E_3 (在线交互) 仅为 2 星, 综合分被拉低至 3.92。TWORAM^[59] 通过将交互轮数降至 2 轮, E_3 提升至 4 星, 综合分提升至 3.98。

(4) 定向防御方案功能覆盖有限: Vaccine^[69] 在 S_3 (抗主动攻击) 上获得 5 星, 但 S_1 、 S_4 、 S_5 均不超过 2 星, 综合分仅 3.62, 说明其定位是特定攻击场景下的补充防御, 而非通用访问模式保护方案。

(5) 混淆路线在效率-安全平衡上最具潜力: OSSE^[37] 和 PANCAKE^[36] 在安全性维度得分均在 3-5 星之间, 同时在 $E1-E4$ 上普遍优于 ORAM 类方案, 代表了当前技术条件下最具工程落地前景的方向。

需要指出的是, 如表 9 所示, 综合评分受权重设置影响显著。当应用场景对可部署性要求更高时 (如 D 维度权重提升至 0.25), PANCAKE 和 SEAL 的排名将上升; 当对完备保护要求更高时 (如 S 维度权重提升至 0.55), Path ORAM 和 OMKX 的排名将显著提升。因此, 其评分应视为特定权重假设下的参考排序, 而非绝对优劣判定。建议研究者在方案选型时, 根据具体应用的安全需求、性能预算和部署约束, 重新调整权重后计算适配自身场景的综合得分。

6 挑战与未来研究方向

基于对现有研究进展与局限性的系统性分析, 本节凝练出可搜索加密访问模式保护领域未来最值得关注的五个核心研究挑战。

6.1 效率-安全极限的逼近探索

当前, ORAM 在统计保护上的性能下界已被证明针对单服务器、同步通信模型下为 $\Omega(\log N)$ 带宽膨胀^[75], 而混淆方案在有限开销下能达到的安全上限仅为统计性保护。两者之间存在一个宽广的“效率-安全间隙”。未来的研究可以尝试从以下方

向突破:

(1) 逼近下界的高效不经意数据结构: 探索新型密码学原语 (如陷门哈希、函数加密、全同态加密的部分增强) 来构造开销更低、更接近理论下限的信息论安全方案。

(2) 自适应保护框架: 现实中的查询并非都要求同等强度的保护。设计根据查询敏感度、数据上下文、系统负载动态调节保护强度的自适应框架, 实现全局能效的最优化。

(3) 多维泄漏统一保护: 将访问模式、搜索模式、体量模式等纳入一个统一的、可量化泄漏的密码学模型, 进而构造能对所有泄漏类型提供一致性和组合保障的集成方案。

6.2 面向新型计算架构的保护方案适配

计算硬件和架构的演进为访问模式保护的性能瓶颈突破提供了新的机遇。

(1) 机密计算的全面赋能: 随着 Intel TDX、AMD SEV-SNP 等新一代 TEE 技术的成熟, EPC 容量和 I/O 性能的约束有望显著缓解。设计深度适配 TEE 特性的专门化访问模式保护方案, 将成为未来数年的热点。

(2) GPU/DPU 加速保护计算: 利用 GPU 的大规模并行计算能力加速 ORAM 路径读取、混淆洗牌、同态加密计算; 利用 DPU (数据处理单元) 在网卡端卸载不经意访问协议, 减少主机 CPU 负载和网络延迟。

(3) 存算一体架构的潜力: 在支持近存储计算的新型硬件上, 探索将部分不经意操作下沉至存储设备内部执行, 从而减少数据搬运并从根本上削弱内存访问模式的物理暴露。

6.3 面向高级搜索功能的访问模式保护

目前绝大多数保护方案仅适用于单关键词精确匹配搜索, 而现实中的数据检索需求日益复杂。

(1) 复杂查询的不经意处理: 设计支持布尔查询、范围查询、时空查询、子图匹配等高级搜索功能的高效访问模式保护方案。这在基因数据检索、金融异常检测等领域具有迫切需求。

(2) 模糊与语义搜索保护: 探索容忍拼写错误、支持同义词扩展的模糊搜索机制与其访问模式保护的结合。如何在隐藏关键词模糊匹配关系的同时, 保证搜索召回率和效率, 是一个待解的难题。

(3) 多模态数据保护: 面向图像、视频、音频

等多模态数据的可搜索加密协议正在兴起, 其特有的特征向量检索和相似度查询在访问模式保护上面临全新建模挑战。

6.4 抗量子计算威胁的后量子保护方案

随着量子计算技术的稳步发展, 现有方案所依赖的 RSA、DH、甚至部分基于格问题的构造, 都面临被量子计算机攻破的远期风险。

(1) 后量子不经意访问原语: 基于 NIST 标准化后量子密码算法 (如 CRYSTALS-Kyber、CRYSTALS-Dilithium) 或新型格困难问题, 构造后量子安全 ORAM、混淆网络和分布式保护协议。

(2) 量子安全模型重定义: 在量子攻击者可以进行叠加查询的更强敌手模型下, 重新审视和定义访问模式保护的安全目标, 并构造具有量子随机安全 (QPR) 证明的方案。

6.5 面向大语言模型时代的攻防演进

LLM 的爆发正在从根本上改变访问模式攻防的假设和手段^[30-31]。

(1) LLM 辅助的智能防御: 利用生成式 AI 实时生成逼真的混淆查询或虚假文档填充, 干扰攻击者对真实访问模式的统计建模。关键在于确保 AI 生成的混淆数据在统计上与用户真实行为不可区分。

(2) 合成数据攻击的形式化防御: 既然 LLM 可以零成本生成合成背景知识来增强攻击, 防御方案就需要在理论上提供针对“任意背景知识 (包括合成知识)”的鲁棒性保证。这将把差分隐私或信息论安全的要求提升到新的高度。

(3) 面向联邦学习与多方协作场景的保护: 在联邦学习等分布式协作计算场景中, 梯度更新、模型参数交换等行为本质上相当于另一种形式的“访问模式泄漏”。将 SSE 场景中的保护思想迁移至这些新兴领域, 是一个极具价值的研究方向。

7 结束语

可搜索加密协议在云计算和大数据时代扮演着“密文可用性”基石的角色。然而, 访问模式泄漏及其引发的系列滥用攻击, 成为制约该技术大规模部署的核心障碍。本文以系统化的方法论, 从泄漏谱系建模、四维技术分类、攻击防御演进、多指标评估、未来挑战前瞻等五个层面, 对可搜索加密的访问模式保护技术进行了全景式的综述。研究表明, 从 ORAM 提供的理论完备保护, 到混淆与差

分隐私实现的高效统计保护,再到分布式架构和硬件 TEE 提供的多样化替代路径,访问模式保护技术已经构建起层次分明、互补有序的解决方案谱系。与此同时,攻击技术从未停滞,从早期的被动统计推断到文件注入再到 LLM 增强攻击,不断挑战着防御的极限。效率与安全之间、保护完备性与部署可行性之间的张力,将持续驱动该领域的创新。展望未来,逼近信息论极限的高效保护方案、面向新型计算架构的硬件-算法协同设计、面向复杂查询与多模态数据的保护泛化、以及在后量子与 LLM 时代的安全再定义,将是这个领域最具挑战性和吸引力的研究方向。我们相信,随着学术研究的深入和产业应用的牵引,可搜索加密的访问模式

保护技术最终将走向成熟,为数据安全与隐私保护提供坚实的技术基座。



龚云平(1992—),男,博士,北京邮电大学网络空间安全专业博士毕业,现为重庆电子科技职业大学专任教师,主要研究方向为大数据安全与隐私保护、密码学和信息安全。Email: gongyunping@bupt.edu.cn



李雄(1984—),男,湖南张家界人,博士,电子科技大学教授、博士生导师,国家级青年人才,爱思唯尔全球前2%顶尖科学家。主要研究方向为数据安全、隐私计算、应用密码学。E-mail: lixiong@uestc.edu.cn



文博(1985—),男,正高,博士生导师南京熊猫电子集团有限公司规划科技部部长,正高级工程师、副总工程师,主要研究方向为边缘计算、数据安全、工业互联网。Email: wenbo@pegcl.cn



武春岭(1975—),男,二级教授,硕士生导师,享受国务院政府特殊津贴专家,重庆市电子信息类学术技术带头人,重庆电子科技职业大学人工智能与大数据学院(网络空间安全学院)院长,主要研究方向为信息安全、密码学。Email: wuch50@126.com



童世华(1980—),男,硕士生导师,三级教授,重庆市电子信息类学术技术带头人后备人选,重庆电子科技职业大学教务处处长,主要研究方向为信息安全、数据隐私保护。Email: 108457135@qq.com

参考文献:

- [1] ZISSIS D, LEKKAS D. Addressing cloud computing security issues[J]. *Future Generation Computer Systems*, 2012, 28(3): 583-592.
- [2] SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]//*Proceedings of the IEEE Symposium on Security and Privacy*. New York: IEEE, 2000: 44-55.
- [3] KAMARA S, PAPAMANTHOU C, ROEDER T. Dynamic searchable symmetric encryption[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2012: 965-976.
- [4] NAVEED M, PRABHAKARAN M, GUNTER C A. Dynamic searchable encryption via blind storage[C]//*Proceedings of the IEEE Symposium on Security and Privacy*. New York: IEEE, 2014: 639-654.
- [5] CASH D, JARECKI S, JUTLA C, et al. Highly-scalable searchable symmetric encryption with support for boolean queries[C]//*Proceedings of the International Cryptology Conference*. Berlin: Springer, 2013: 353-373.
- [6] CASH D, GRUBBS P, PERRY J, et al. Leakage-abuse attacks against searchable encryption[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2015: 668-679.
- [7] SUN S, YUAN X, LIU J K, et al. Practical backward-secure searchable encryption from symmetric puncturable encryption[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2018: 763-780.
- [8] WANG C, CAO N, LI J, et al. Secure ranked keyword search over encrypted cloud data[C]//*Proceedings of the IEEE International Conference on Distributed Computing Systems*. New York: IEEE, 2010: 253-262.
- [9] LI J, WANG Q, WANG C, et al. Fuzzy keyword search over encrypted data in cloud computing[C]//*Proceedings of the IEEE International Conference on Computer Communications*. New York: IEEE, 2010: 1-5.
- [10] CURTMOLA R, GARAY J, KAMARA S, et al. Searchable symmetric encryption: Improved definitions and efficient constructions[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2006: 79-88.
- [11] 李经纬, 贾春福, 刘哲理, 等. 可搜索加密协议研究综述[J]. *软件学报*, 2015, 26(1): 109-128.
LI J W, JIA C F, LIU Z L, et al. Survey of searchable encryption[J]. *Journal of Software*, 2015, 26(1): 109-128.
- [12] ISLAM M S, KUZU M, KANTARCIOGLU M. Access pattern disclosure on searchable encryption: Ramification, attack and mitigation[C]//*Proceedings of the Network and Distributed System Security Symposium*. Reston: Internet Society, 2012: 1-12.
- [13] LIU C, ZHU L, WANG M, et al. Search pattern leakage in searchable encryption: Attacks and new construction[J]. *Information Sciences*, 2014, 265: 176-188.
- [14] ZHANG Y, KATZ J, PAPAMANTHOU C. All your queries are belong to us: The power of file-injection attacks on searchable encryption[C]//*Proceedings of the USENIX Security Symposium*. Berkeley: USENIX Association, 2016: 707-720.
- [15] GRUBBS P, MCPHERSON R, NAVEED M, et al. Breaking web applications built on top of encrypted data[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2016: 1353-1364.
- [16] POULIOT D, WRIGHT C V. The shadow nemesis: Inference attacks on efficiently deployable, efficiently searchable encryption[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2016: 1341-1352.
- [17] OYA S, KERSCHBAUM F. Hiding the access pattern is not enough: Exploiting search pattern leakage in searchable encryption[C]//*Proceedings of the USENIX Security Symposium*. Berkeley: USENIX Association, 2021: 127-142.
- [18] MARKATOU E A, TAMASSIA R. Full database reconstruction with access and search pattern leakage[C]//*Proceedings of the International Conference on Information Security*. Berlin: Springer, 2019: 25-43.
- [19] LAMBREGTS S, CHEN H, NING J, et al. VAL: Volume and access pattern leakage-abuse attack with leaked documents[C]//*Proceedings of the European Symposium on Research in Computer Security*. Berlin: Springer, 2022: 653-676.
- [20] GRUBBS P, SEKNIQI K, BINDSCHAEDLER V, et al. Leakage-abuse attacks against order-revealing encryption[C]//*Proceedings of the IEEE Symposium on Security and Privacy*. New York: IEEE, 2017: 655-672.
- [21] POH G S, CHIN J J, YAU W C, et al. Searchable symmetric encryption: Designs and challenges[J]. *ACM Computing Surveys*, 2017, 50(3): 1-37.
- [22] DU M, WANG Q, HE M, et al. Privacy-preserving indexing and query processing for secure dynamic cloud storage[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(9): 2320-2332.
- [23] 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.
SHEN Z R, XUE W, SHU J W. Research and progress of searchable encryption[J]. *Journal of Software*, 2014, 25(4): 880-895.
- [24] FULLER B, VARIA M, YERUKHIMOVICH A, et al. SoK: Cryptographically protected database search[C]//*Proceedings of the IEEE Symposium on Security and Privacy*. New York: IEEE, 2017: 172-191.
- [25] 赵陈斌, 李金美, 何琨, 等. 对称可搜索加密研究综述[J]. *密码学报*, 2026, 13(1): 1-27.
ZHAO C B, LI J M, HE K, et al. Survey of symmetric searchable encryption[J]. *Journal of Cryptologic Research*, 2026, 13(1): 1-27.
- [26] STEFANOV E, VAN DIJK M, SHI E, et al. Path ORAM: An extremely simple oblivious RAM protocol[C]//*Proceedings of the ACM Conference on Computer and Communications Security*. New York: ACM Press, 2013: 299-310.
- [27] CHEN G, LAI T H, REITER M K, et al. Differentially private access patterns for searchable symmetric encryption[C]//*Proceedings of the IEEE International Conference on Computer Communications*. New York: IEEE, 2018: 810-818.
- [28] DEMERTZIS I, PAPADOPOULOS D, PAPAMANTHOU C, et al. SEAL: Attack mitigation for encrypted databases via adjustable leakage[C]//*Proceedings of the USENIX Security Symposium*. Berkeley: USENIX Association, 2020: 2433-2450.
- [29] VO V, LAI S, YUAN X, et al. Accelerating forward and backward private searchable encryption using trusted execution[C]//*Proceedings of the International Conference on Applied Cryptography and Network Security*. Berlin: Springer, 2020: 83-103.
- [30] CHIU J, PAUL P P, WAHAB Z. Enhancing leakage attacks on searchable symmetric encryption using LLM-based synthetic data generation

- [J]. Cryptology ePrint Archive, 2025, 548: 1-18.
- [31] KHAN S, KHAN M, KHAN M A, et al. A blockchain-enabled AI-driven secure searchable encryption framework for medical IoT systems[J]. IEEE Journal of Biomedical and Health Informatics, 2025.
- [32] FABER S, JARECKI S, KRAWCZYK H, et al. Rich queries on encrypted data: Beyond exact matches[C]//Proceedings of the European Symposium on Research in Computer Security. Berlin: Springer, 2015: 123-145.
- [33] KAMARA S, MOATAZ T. Computationally volume-hiding structured encryption[C]//Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2019: 183-213.
- [34] GEORGE M, KAMARA S, MOATAZ T. Structured encryption and dynamic leakage suppression[C]//Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2021: 370-396.
- [35] PATEL S, PERSIANO G, YEO K. What storage access privacy is achievable with small overhead?[C]//Proceedings of the ACM Symposium on Principles of Database Systems. New York: ACM Press, 2019: 182-199.
- [36] GRUBBS P, KHANDLWAL A, LACHARITÉ M S, et al. Pancake: Frequency smoothing for encrypted data stores[C]//Proceedings of the USENIX Security Symposium. Berkeley: USENIX Association, 2020: 2451-2468.
- [37] SHANG Z, OYA S, PETER A, et al. Obfuscated access and search patterns in searchable encryption[C]//Proceedings of the Network and Distributed System Security Symposium. Reston: Internet Society, 2021.
- [38] DUAN S, LIU H, GU L. A searchable encryption scheme with hidden search pattern and access pattern on distributed cloud system[J]. Peer-to-Peer Networking and Applications, 2023, 16: 1716-1738.
- [39] HOANG T, YAVUZ A A, GUJARDO J. A secure searchable encryption framework for privacy-critical cloud storage services[J]. IEEE Transactions on Services Computing, 2021, 14(6): 1675-1689.
- [40] HOANG T, YAVUZ A A, DURAK F B, et al. A multi-server oblivious dynamic searchable encryption framework[J]. Journal of Computer Security, 2019, 27(6): 649-676.
- [41] FUHR B, BAHMANI R, BRASSER F, et al. HardIDX: Practical and secure index with SGX[C]//Proceedings of the IFIP Annual Conference on Data and Applications Security and Privacy. Berlin: Springer, 2017: 386-408.
- [42] SHAON F, KANTARCIOGLU M, LIN Z, et al. SGX-MR-Prot: Efficient and developer-friendly access-pattern protection in trusted execution environments[C]//Proceedings of the IEEE International Conference on Distributed Computing Systems. New York: IEEE, 2023: 1029-1032.
- [43] 张玉清, 王晓菲, 刘雪峰, 等. 云计算环境安全综述[J]. 软件学报, 2016, 27(6): 1328-1348.
- ZHANG Y Q, WANG X F, LIU X F, et al. Survey on cloud computing security[J]. Journal of Software, 2016, 27(6): 1328-1348.
- [44] 刘源龙, 戴华, 李张晨, 等. 云环境中语义感知密文检索研究综述[J]. 计算机科学, 2024, 51(11): 298-306.
- LIU Y L, DAI H, LI Z C, et al. Survey of semantic-aware ciphertext retrieval in cloud environments[J]. Computer Science, 2024, 51(11): 298-306.
- [45] 严莉, 殷田, 刘培顺, 等. 基于属性的可搜索加密综述[J]. 计算机科
- 学, 2024, 51(S2): 842-853.
- YAN L, YIN T, LIU P S, et al. Survey of attribute-based searchable encryption[J]. Computer Science, 2024, 51(S2): 842-853.
- [46] 吴鹏飞, 沈晴霓, 秦嘉, 等. 不经意随机访问机研究综述[J]. 软件学报, 2018, 29(9): 2753-2777.
- WU P F, SHEN Q N, QIN J, et al. Survey of oblivious random access machine[J]. Journal of Software, 2018, 29(9): 2753-2777.
- [47] CHASE M, KAMARA S. Structured encryption and controlled disclosure[C]//Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security. Berlin: Springer, 2010: 577-594.
- [48] ABDALLA M, BELLARE M, CATALANO D, et al. Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions[J]. Journal of Cryptology, 2008, 21(3): 350-391.
- [49] BAEK J, SAFARI-NAINI R, SUSILO W. Public key encryption with keyword search revisited[C]//Proceedings of the International Conference on Computational Science and Its Applications. Berlin: Springer, 2008: 1249-1259.
- [50] BOST R, MINAUD B, OHRIMENKO O. Forward and backward private searchable encryption from constrained cryptographic primitives [C]//Proceedings of the ACM Conference on Computer and Communications Security. New York: ACM Press, 2017: 1465-1482.
- [51] ZHANG X, WANG W, XU P, et al. High recovery with fewer injections: Practical binary volumetric injection attacks against dynamic searchable encryption[C]//Proceedings of the USENIX Security Symposium. Berkeley: USENIX Association, 2023: 5953-5970.
- [52] ZHANG L, WANG J, WU J, et al. Violin: Powerful volumetric injection attack against searchable encryption with optimal injection size[J]. IEEE Transactions on Dependable and Secure Computing, 2025, 22(4): 4103-4115.
- [53] CHEN T, XU P, WANG W, et al. Bestie: Very practical searchable encryption with forward and backward security[C]//Proceedings of the European Symposium on Research in Computer Security. Berlin: Springer, 2021: 3-23.
- [54] BOST R. Σ ορος: Forward secure searchable encryption[C]//Proceedings of the ACM Conference on Computer and Communications Security. New York: ACM Press, 2016: 1143-1154.
- [55] GHAREH CHAMANI J, PAPADOPOULOS D, PAPAMANTHOU C, et al. New constructions for forward and backward private symmetric searchable encryption[C]//Proceedings of the ACM Conference on Computer and Communications Security. New York: ACM Press, 2018: 1038-1055.
- [56] POLICICCHIO S M, YAVUZ A A. Preventing memory access pattern leakage in searchable encryption[C]//Proceedings of the iConference. Berlin: Springer, 2015.
- [57] VAN BULCK J, MINKIN M, WEISSE O, et al. Foreshadow: Extracting the keys to the Intel SGX kingdom with transient out-of-order execution[C]//Proceedings of the USENIX Security Symposium. Berkeley: USENIX Association, 2018: 991-1008.
- [58] GRUBBS P, LACHARITÉ M S, MINAUD B, et al. Pump up the volume: Practical database reconstruction from volume leakage on range queries[C]//Proceedings of the ACM Conference on Computer and Communications Security. New York: ACM Press, 2018: 315-331.
- [59] GOLDBREICH O, OSTROVSKY R. Software protection and simulation on oblivious RAMs[J]. Journal of the ACM, 1996, 43(3): 431-473.

- [60] GARG S, MOHASSEL P, PAPAMANTHOU C. TWORAM: Efficient oblivious RAM in two rounds with applications to searchable encryption[C]//Proceedings of the International Cryptology Conference. Berlin: Springer, 2016: 563-592.
- [61] DEMERTZIS I, PAPAMANTHOU C, TALKAD V. Searchable encryption with optimal locality: Achieving sublogarithmic read efficiency [C]//Proceedings of the International Cryptology Conference. Berlin: Springer, 2018: 371-406.
- [62] DEGITZ A, HARTENSTEIN H. PATCONFDB: Design and evaluation of access pattern confidentiality-preserving indexes[J]. Transactions on Data Privacy, 2018, 11(2): 81-109.
- [63] REN K, WANG C, WANG Q. Security challenges for the public cloud [J]. IEEE Internet Computing, 2012, 16(1): 69-73.
- [64] DWORK C. Differential privacy[C]//Proceedings of the International Colloquium on Automata, Languages and Programming. Berlin: Springer, 2006: 1-12.
- [65] ISHAI Y, KUSHILEVITZ E, LU S, et al. Private large-scale databases with distributed searchable symmetric encryption[C]//Proceedings of the Cryptographers' Track at the RSA Conference. Berlin: Springer, 2016: 90-107.
- [66] DAUTERMAN E, FENG E, LUO E, et al. DORY: An encrypted search system with distributed trust[C]//Proceedings of the USENIX Symposium on Operating Systems Design and Implementation. Berkeley: USENIX Association, 2020: 1101-1119.
- [67] HOANG T, YAVUZ A A, GUAJARDO J. Practical and secure dynamic searchable encryption via oblivious access on distributed data structure[C]//Proceedings of the Annual Computer Security Applications Conference. New York: ACM Press, 2016: 302-313.
- [68] LI S, JING X, WANG Y, et al. Practical searchable encryption scheme against response identity attacks[J]. Information Sciences, 2025, 706: 121975.
- [69] LIU H, WANG B, NIU N, et al. Vaccine: Obfuscating access pattern against file-injection attacks[C]//Proceedings of the IEEE Conference on Communications and Network Security. New York: IEEE, 2019: 1-9.
- [70] JIANG Q, QI Y, QI S, et al. Pbsx: A practical private boolean search using Intel SGX[J]. Information Sciences, 2020, 521: 174-194.
- [71] MCKEEN F, ALEXANDROVICH I, BERENZON A, et al. Innovative instructions and software model for isolated execution[C]//Proceedings of the Workshop on Hardware and Architectural Support for Security and Privacy. New York: ACM Press, 2013: 1-8.
- [72] YANG Y, HOU T, FAN H, et al. A practical dynamic searchable symmetric encryption scheme with trusted hardware-assisted[J]. IEEE Transactions on Computers, 2026.
- [73] HOANG T, OZMEN M O, JANG Y, et al. Hardware-supported ORAM in effect: Practical oblivious search and update on very large dataset[J]. Proceedings on Privacy Enhancing Technologies, 2019, (1): 172-191.
- [74] NING J, HUANG X, POH G S, et al. LEAP: Leakage-abuse attack on efficiently deployable, efficiently searchable encryption with partially known dataset[C]//Proceedings of the ACM Conference on Computer and Communications Security. New York: ACM Press, 2021: 2307-2320.
- [75] LARSEN K G, NIELSEN J B. Yes, there is an oblivious RAM lower bound! [C]//Proceedings of the International Cryptology Conference. Berlin: Springer, 2018: 523-542.
- [76] 张文琪, 李雄, 尹智明, 等. 鲁棒的前后向隐私联合对称可搜索加密方案[J]. 软件学报, 2025, 36(8): 3858-3882.
- ZHANG W Q, LI X, YIN Z M, et al. Robust forward and backward private conjunctive symmetric searchable encryption scheme[J]. Journal of Software, 2025, 36(8): 3858-3882.
- [77] 聂旭云, 成驰, 耿聪, 等. 支持模式隐藏的多关键词公钥可搜索加密方案[J]. 通信学报, 2025, 46(3): 131-143.
- NIE X Y, CHENG C, GENG C, et al. Multi-keyword public key searchable encryption scheme with pattern hiding[J]. Journal on Communications, 2025, 46(3): 131-143.