利用 eBPF 监控改进可搜索对称加密中的泄漏滥用攻击

Chinecherem Dimobi弗
吉尼亚理工学院
chinedimobi@vt.edu

arXiv:2603.07030v1 [cs.CR] 7 Mar 2026

摘要—**可搜索对称加密（SSE）允许用户在存储于不可信服务器（如云服务提供商）上的加密数据中进行搜索。虽然 SSE 会隐藏查询和文档的内容，但仍会泄露某些模式，例如某个查询被执行的频率。研究表明，这些泄露可被用于实施泄漏滥用攻击，但近期的防御措施已使此类攻击更难得逞。在本研究中，我们探讨如何利用 eBPF（扩展伯克利包过滤器）进行系统级监控，以揭示超出 SSE 威胁模型通常涵盖范围的新型泄露。通过观察搜索操作期间的底层系统行为，我们展示了攻击者如何获取有关查询行为、文档访问和处理流程的更多信息。我们根据这些观察定义了一种新的泄漏模式，并展示了它们如何增强现有攻击。研究结果表明，系统级泄露对 SSE部署构成了现实威胁，在设计防御措施时必须予以考虑。本研究为弥合理想化的可搜索对称加密安全性与系统级暴露的现实情况之间的差距迈出了一步。**

索引词——**可搜索对称加密、eBPF、泄漏滥用攻击**

一、引言

可搜索对称加密（SSE）使用户能够搜索存储在外部服务器上的加密数据 [1]。这已成为保护外包存储隐私的重要工具，尤其是在使用云服务时。其基本原理是，用户可以上传加密文件，同时执行关键词搜索，而不泄露文件内容或查询内容。然而，可搜索对称加密系统提供的安全保障仍然有限。已知这类系统会泄露某些信息，例如某个查询被执行的次数，或其匹配的文档数量。这些泄露可能会被利用，实施所谓的泄漏滥用攻击。

一种常见的泄漏滥用攻击是频率匹配攻击（FMA） [2], ，攻击者通过观察搜索查询的频率，并将其与已知关键词分布关联起来，从而推断搜索词。不过，研究人员在防御此类攻击方面已取得进展 [3], ，提出了填充查询结果或隐藏数据量等策略。这些改进使攻击者更难仅凭查询模式获取大量信息。

本项目转换思路，提出了一个不同的问题：如果攻击者不仅能访问查询模式，还能观察系统本身的底层行为，会发生什么？具体来说，如果攻击者能够监控搜索操作期间内核的活动，会怎样？这正是 eBPF（扩展伯克利数据包过滤器）发挥作用的地方。eBPF [4], [5] 是一种 Linux内核扩展，可让程序在内核中运行，并实时观察或响应系统事件。它常用于调试和性能监控，但也可用于跟踪诸如哪些文件被打开以及打开顺序等信息。

本研究旨在探讨通过 eBPF 进行系统级监控，能否揭示现有防御措施未曾考虑到的可搜索对称加密系统中的新泄露。为此，我们力求：

- 调查可搜索对称加密搜索操作期间发生了哪些底层事件。
- 识别可从这些事件中观察到的其他泄漏模式。
- 展示如何利用这些信息来增强或构造更有效的泄露滥用攻击。

本研究的贡献如下：

- 我们识别并定义了一种新的泄漏模式 $L_{fileAccess}$，它可以揭示查询期间访问的确切文件。
- 我们展示了如何利用这种泄露更准确地恢复查询，即使传统的频率攻击失效时也能做到。
- 我们强调了可搜索对称加密理论安全模型与现实部署漏洞之间的差距，并指出了未来防御措施的发展方向。

本研究旨在推动人们以更广泛的视角看待可搜索对称加密安全性，将搜索操作运行时系统层面可观察到的信息也纳入考量。

II. 相关工作

A. 可搜索对称加密方案的演进

可搜索对称加密最初由 Song 等人提出，他们设计了一种可直接在加密数据上执行关键词搜索的方法。其研究证明，在实现某种形式的搜索功能的同时保护隐私是可行的。然而，他们的方案存在局限：需要逐一扫描每个文件，且未明确界定哪些信息可能会被无意泄露。

Curtmola 等人 [1] 后来提出了一种更为结构化的可搜索对称加密方法。他们引入了清晰的安全定义，并形式化了如何对泄露进行推理。他们的方案将加密索引与加密数据分离，使搜索更快、更易于管理。这项工作成为许多后续设计的基线。

存储系统迁移到云端，现实世界的数据集变得更加动态，由此催生了动态可搜索对称加密，它支持添加或删除文档等更新。Kamara 等人 [7] 提出了最早的动态可搜索对称加密方案之一。他们的工作允许在保持安全性的同时更新加密数据，但也引入了新的泄漏向量。

后续研究着重于提升可搜索对称加密方案的表达能力和可扩展性，使其支持更丰富的查询、多客户端环境以及大规模场景下的性能 [8]。然而，随着可搜索对称加密系统功能日益强大，管理和理解泄露的复杂性也随之增加，从而导致功能与隐私之间需要权衡。

B. 可搜索对称加密系统中的泄漏

搜索操作中的泄露主要表现为访问模式（揭示针对给定查询返回了哪些文档）和搜索模式（表明特定查询是否被重复执行）。

Islam 等人 [9] 指出，随着时间推移观察访问模式，并结合有关数据的背景知识，就可能还原原始查询。Oya 和 Kerschbaum [2]进一步指出，仅隐藏访问模式并不足够，因为重复查询仍可能泄露信息。近期，Xu 等人 [10] 指出，即使是前向与后向隐私系统，也会通过数据量和共现泄露数据。这些细微的泄露，例如文档共同出现的频率，仍可能帮助攻击者重建查询意图。

C. 泄漏滥用攻击

泄漏滥用攻击会利用这些泄漏模式，但不会破解底层加密。Cash 等人 [11]介绍了利用频率分析将加密查询映射到已知关键词分布的攻击。

Gui 等人 [12] 研究了体积泄露，攻击者通过返回的文档数量推断查询意图。Bakas 和 Michalas [13] 展示了针对多客户端可搜索对称加密系统的攻击，其中共享索引结构会导致跨客户端泄露。

这些研究表明，如果攻击者能够利用系统在正常运行过程中泄露的信息，仅有加密保护是不够的。这促使人们制定更广泛的威胁模型，将可搜索对称加密与其环境之间的相互作用纳入考量。

D. 使用 *eBPF* 进行系统级监控

eBPF（扩展伯克利数据包过滤器） [14] 是一项 Linux 内核技术，最初用于数据包过滤。如今，它已发展为一个强大的框架，可用于观察和响应系统事件 [15], [16]。它允许小型程序在内核中安全运行，从而实现低开销跟踪、系统调用过滤 [4]、容器安全 [17], 和网络监控[18]。

Cilium 和 Tracee 等工具利用 eBPF 提供深入的运行时可见性。由于 eBPF 在内核中运行并受到严格的安全检查，因此能够捕获详细的系统行为，同时不影响性能 [19], [20]。

Lu 等人 [21] 和 Nguyen 等人 [22] 的最新研究表明，文件访问和输入/输出操作等系统级模式会造成密码学威胁模型未涵盖的泄露。

我们的研究建立在这一思路之上。我们探讨攻击者是否能够借助eBPF等工具所提供的系统级可见性，在可搜索对称加密操作过程中推断敏感模式。例如，观察文件访问调用可能揭示加密查询如何与数据交互，从而形成现有模型尚未考虑的一类新泄露。我们的研究结果推动人们以更整体的视角看待可搜索对称加密安全性，将系统本身在加密搜索过程中所暴露的信息纳入考量。

III. 预备知识

A. 可搜索对称加密的工作原理：技术概述

对称可搜索加密（可搜索对称加密）方案旨在支持对存储于云服务器等不可信平台上的加密文档进行关键词搜索。其核心思想是，云端应能够响应搜索查询，同时不获知文档内容、被搜索的关键词或返回的结果。

1) 倒排索引构建：该过程从数据所有者一侧开始。所有文档都会被解析以提取关键词。随后，这些关键词用于构建倒排索引：一种将每个关键词映射到该关键词出现位置的文档标识符列表（例如文件ID）的数据结构。例如：

"发票" → {doc1, doc3,doc7} "合同" → {doc2, doc4}

该索引至关重要，因为它支持高效检索，无需每次搜索都扫描所有加密文件。

2）索引和文档加密：每份文档都使用对称加密方案进行加密。同时，倒排索引采用确定性加密或结构化加密进行加密，使搜索令牌能够与索引条目匹配，而不会泄露明文内容。随后，加密索引和加密文档会上传至云服务提供商（CSP）。CSP 存储这些内容，但无法对其进行解密。

3）搜索协议：执行搜索时，客户端（数据所有者或授权用户）使用密钥生成与目标关键词对应的确定性搜索令牌。该令牌（通常记作 t_w ，对应关键词 w）会发送给 CSP。

CSP 使用 t_w 搜索加密索引，并查找与该关键词关联的加密文档标识符。找到后，CSP 会返回与这些标识符对应的加密文件。

然后，客户端使用其密钥在本地解密文档。关键词或解密后的文档绝不会泄露给云服务提供商。

4) 更新协议（添加/删除）：动态可搜索对称加密支持对加密数据库进行更新，包括添加和删除文档。

- **添加：**添加新文档时，客户端会解析该文档以提取关键词，更新本地

倒排索引，对文档和新的索引条目进行加密，并将它们发送给云服务提供商。

- **删除：**当文档被删除时，客户端向云服务提供商发送删除令牌，后者移除相应的加密文档并更新受影响的索引条目。

为了维持安全性，更新操作同样需要加密协议，以防止云服务提供商获知具体发生了哪些变化，除受影响条目的数量之外。

5）泄露类型：每种可搜索对称加密操作都会泄露略有不同的一类元数据：

- **添加**（L_{add}）会泄露所添加新文档中的关键词数量。
- **更新**（L_{update}）会泄露被修改或添加的索引条目数量。
- **删除**（L_{delete}）会泄露被移除的索引条目数量。
- **搜索**（L_{search}）会泄露针对给定关键词返回的文档数量。

这类泄露大多源于索引中的加密标识符会长期保持可关联性，即使其内容已被隐藏。

B. 频率匹配攻击（*FMA*）

频率匹配攻击（FMA）是一种经典的泄露滥用技术，它利用查询和响应中的重复结构。攻击者首先观察一系列搜索令牌，并记录每个令牌的使用频率（即查询频率），以及每个令牌检索到的文档数量（即结果大小）。由此可得到加密查询的频率分布。

如果攻击者可以访问辅助数据集（例如类似的明文数据集或先验知识），就能计算该数据集中已知关键词的频率分布。通过匹配这两个分布，攻击者试图恢复加密搜索令牌与明文关键词之间的映射。

当攻击者能够观察搜索的结果大小（L_{search}）时，FMA 的成功率会显著提高。结果大小独特的关键词通常可以与特定搜索令牌一一匹配，即使其频率存在歧义。

然而，FMA 并非完美。其主要局限之一是多个关键词具有相似的频率分布。例如，如果“发票”“合同”和“预算”各自出现在 20% 的文档中，那么它们的令牌就会具有相似的频率和结果大小特征。在这种情况下，由于不存在唯一的映射，FMA 无法有把握地将令牌与关键词对应起来。

这一局限为新的向量留下了可能性。如果频率信息不足，是否还有其他信息，例如系统级信息，能够揭示 FMA 无法揭示的内容？这个问题正是我们工作的出发点。

C. eBPF 的工作原理

内核是任何系统的核心组件，负责管理该系统的所有资源。扩展伯克利包过滤器（eBPF）是一种 Linux 内核技术，无需修改源代码或加载自定义内核模块，即可直接在内核中运行小型沙箱程序。eBPF 程序能够实时、安全地与各种内核组件交互，从而深入了解系统行为和性能。

1) eBPF 的核心概念：eBPF 通过将小型程序附加到特定的内核事件上来运行。这些事件可能包括系统调用（如 open、read 或 write）、网络事件（如传入的数据包），甚至包括在内核中定义的自定义跟踪点。

每个 eBPF 程序在执行前，都会由一个名为 *eBPF* 验证器的内置安全检查器进行验证。该验证器确保 eBPF 程序不会导致内核崩溃、能在有限数量的指令内完成执行，并且只访问获准访问的内核内存。通过验证后，程序会在内核事件触发时运行。事件发生时，eBPF 程序便会执行，并可能记录或分析相关信息。

IV. 威胁模型与安全假设

A. 对手模型

我们假设采用**诚实但好奇**的对手模型。换言之，云服务提供商会忠实地执行所有协议，不会恶意篡改存储的数据或查询。不过，云服务提供商具有好奇心，会试图从处理查询和管理数据时自然泄露的信息中尽可能多地获取信息。

具体而言，对手具备以下能力：

1) **全面观察加密操作：**攻击者可以看到云服务提供商处理的所有加密文档、加密索引和查询令牌。2) **系统级监控权限：**借助 eBPF 等工具，攻击者可以观察云服务提供商系统的详细行为，例如每次操作期间访问了哪些加密文件、系统调用的顺序，甚至可能获取时序信息。

3) **辅助（背景）知识：**攻击者对底层数据集有部分了解，例如关键词频率分布或数据统计信息。这些知识可能来自类似的明文数据集或公开来源。对手对明文索引也有部分了解，包括知道某些文档分布。

然而，对手无法直接破解底层密码学原语。所使用的加密方案是安全的，密钥由数据所有者保密。

B. 安全假设

为了清晰地界定攻击者的视角，我们基于以下假设：

- 所有关键词、文档和索引均使用安全、成熟的密码学方法进行加密。对手在没有密钥的情况下无法直接解密或逆转加密方案。
- 加密索引中的文档标识符被视为不透明信息。在传统的可搜索对称加密模型中，这些 ID 不应泄露其所代表文档的有意义信息。
- 数据所有者与云服务提供商之间的通信受到保护。攻击者只能看到由此产生的加密查询和响应，而看不到明文通信信道。

C. 攻击目标

对手的主要目标是恢复客户端发出的明文搜索查询。

我们尤其关注泄漏滥用攻击，例如频率匹配攻击（FMA）。这类攻击会利用元数据泄漏，例如每次查询返回的文档数量。

D. 范围与局限性

在我们的威胁模型中，我们不考虑能够主动修改查询或已存储数据的攻击者。我们的模型仅关注被动攻击，这与现实的云场景相符：云服务提供商会正确执行协议，但会尝试通过被动方式推断更多信息。

V. 方法

本节介绍我们采用的方法，用于研究系统级监控带来的额外泄露是否会显著增强针对可搜索对称加密的泄漏滥用攻击，尤其关注 eBPF 监控如何帮助攻击者更准确地推断查询。

A. 直觉与高层思路

我们的直觉基于这样一个事实：标准的可搜索对称加密方案往往忽略了云服务提供商在系统级与加密数据交互时产生的泄露。例如，当云服务提供商处理搜索查询时，必须访问存储中的某些加密文件，并将其返回给客户端。尽管这些文件和索引都经过加密，但文件名本身通常不会改变。这意味着，在操作系统层面可见的文件名与原始明文文件名直接对应。

我们利用 eBPF 准确观察每次搜索期间访问了哪些加密文件。由于 eBPF 能够安全地监控内核事件，因此我们可以轻松跟踪查询期间访问的密文文件的文件名。鉴于我们假设攻击者对明文数据集和索引具有部分知识，直接观察这些文件名便提供了一个此前未被充分考虑的新泄漏向量。该方法使攻击者能够将加密查询中的文件名与已知的明文文件名进行匹配，从而揭示加密令牌与明文关键词之间的关系。

B. 定义的泄漏向量 $L_{fileAccess}$

我们的方法引入了一种新的系统级泄漏模式，称为 $L_{fileAccess}$。$L_{fileAccess}$ 背后的核心理念简单却十分强大：

1) 每当客户端向云服务提供商发送搜索查询时，云服务提供商都会通过查找加密索引并访问特定的密文文件来处理该查询。2) 通过在关键系统调用（例如 openat、read）处直接向内核附加 eBPF 监控程序，我们可以准确记录每个查询访问了哪些密文文件。

形式上，我们将新的泄漏向量 $L_{fileAccess}$ 定义如下：

对于每个搜索令牌 t_w，泄露会揭示以下集合：

$$t_w \to \{f_1, f_2, \ldots, f_n\}$$

此处， f_i 表示作为查询 t_w的一部分被访问并返回给客户端的密文文档的文件名。

这种新的泄漏向量补充了结果大小和频率模式等传统泄露信息。攻击者不再只是知道某个查询返回了多少个文件，而是能够准确知道每次搜索对应哪些密文文件。

C. 增强型频率匹配攻击（*eFMA*）

为了实际利用这一新的泄露向量（$L_{fileAccess}$），我们将标准的频率匹配攻击（FMA）扩展为所谓的 eBPF 增强型频率匹配攻击（eFMA）。eFMA 的目标很明确，即恢复每个加密搜索查询背后的明文关键词。eFMA 的流程如下：

- 首先，观察每个查询令牌，并记录两项关键信息：1）每个查询的 结果长度（返回的文件数量）；2）该查询访问的加密文档的完整集合，这些信息通过 eBPF 监控获取。
- 接下来，为每个查询令牌构建候选映射。具体而言，将每个查询中观测到的密文文件（F_e）与根据部分知识已知的明文文件名（F_p）进行比较。
- 我们首先进行频率分析。具有唯一频率（不同结果长度）的查询可以直接与明文关键词匹配。然而，当多个关键词具有相似频率时，传统的 FMA 难以将它们区分开来。
- 当频率相同时，我们利用查询期间访问的确切密文文件名集合。如果密文文件名与明文索引中的文件名（$F_e = F_p$）完全匹配，我们便可直接推断令牌对应的关键词。
- 形式上，该映射可表示为：

$$t_w = w \quad \text{if} \quad F_e = F_p$$

其中 t_w 是关键词 w的加密令牌，而F_e、 F_p 分别是密文文件集和明文文件集。

与仅依靠传统的基于频率的攻击相比，这种增强版匹配方法极大地提高了我们恢复查询的能力。

D. 协议分析

我们的协议利用 eBPF 所实现的实际观测，显著强化了传统的泄漏滥用攻击。此处的一个关键洞见是，可搜索对称加密方案通常不会在文件系统层面重命名或混淆文件名。这一微小疏忽形成了一条强大的泄漏通道。

从对抗者的角度来看，这意味着即便是前向隐私或理论上安全的可搜索对称加密方案，仍然容易受到文件访问模式等系统级泄漏的影响。显然，安全模型需要明确考虑此类泄露，而不能仅限于密码学保证或标准的索引级泄漏。

总之，通过借助 eBPF 集成系统级文件访问监控，我们的增强版方法显著推进了泄漏滥用攻击，凸显了可搜索对称加密安全性中一个重要且常被忽视的维度。

VI. 实现与实验

A. 系统设置与配置

我们的实现采用具有前向隐私的 DK-Nguyen动态可搜索对称加密方案 [23] ，该方案遵循标准的三方模型，涉及数据所有者、可信机构和云服务提供商（CSP）。数据所有者对一组文档进行加密，并构建相应的加密索引。随后，加密数据会上传至云服务提供商，之后在此执行查询。

我们在虚拟机上的 Docker 容器中运行动态可搜索对称加密方案，并使用 Enron [24]数据集中的 100 封邮件作为实验语料。选择该数据集是因为它贴近现实且规模适中，适合观察查询级泄漏。为专注于检索过程中产生的泄漏，我们将评估范围限定为仅搜索查询（不进行更新）。

为了追踪这些查询期间的系统级活动，我们使用了 bpftrace [25], ，这是一种基于 eBPF 构建的动态追踪工具。bpftrace 使我们能够监控文件级系统调用，尤其是文档访问期间触发的调用，而无需修改动态可搜索对称加密代码库。

B. 评估指标

我们在评估中使用的主要指标是 **查询恢复准确率**，用于衡量攻击者将加密搜索令牌映射到对应明文关键词的准确程度。该指标适用于我们的研究，因为基线攻击和增强攻击的核心目标都是根据可观测泄漏推断明文查询。

我们比较了两种攻击：

- **基线频率匹配攻击（FMA）：** 仅使用每个查询的结果大小（返回的文档数量）来执行基于频率的推断。
- **增强型频率匹配攻击（eFMA）：** 同时使用结果大小和每次查询期间访问的密文文件集合，这些信息通过 eBPF 监控获得。

C. 实验设计

为模拟对手，我们采用了被动攻击设置。我们充当云服务提供商，观察查询令牌以及每次查询返回的文件，同时不修改或干扰 DSSE 代码。

a) 关键词选择：我们从 Enron数据集中查询了一组经过筛选的关键词，确保其中既有高频词，也有低频词。由此形成了一系列频率分布，既展现了仅基于频率的攻击的优势，也暴露了其弱点。

b) 泄露收集：

- 对于基线攻击，我们收集了每次查询的结果大小。
- 对于增强攻击，我们使用 bpftrace 记录了每次查询期间访问的具体密文文件名。这些文件名在存储系统中未经过混淆处理，因此我们能够观察到精确的文档访问模式。

c) 攻击者知识：我们假设攻击者可以部分访问明文数据集（例如，来自 Enron 的公开副本），并且知道关键词与其出现文档之间的映射。这些知识使攻击者能够构建候选倒排索引，用于频率匹配和文档集合比较。

D. 结果与分析

我们的实验结果汇总于表 I 和图 1。

关键词恢复准确率对比（表 I）

方法	查询恢复准确率
基线（FMA）	77.8%
增强版（eFMA）	100%

基线 FMA 的查询恢复准确率达到 77.8%。在通过 eBPF 跟踪加入系统级泄漏后，增强攻击的准确率达到 100%。我们的实验结果清楚地表明，借助 eBPF 捕获的文件级泄漏如何提高查询恢复的准确率。图 1 展示了在基线频率匹配攻击（FMA）与我们的增强版 eFMA 方法下，每个令牌的映射结果明细。

每个点代表一个令牌。蓝色点表示基线结果，橙色点表示加入基于 eBPF 的文件访问泄露后的结果。我们可以看到，像 **T12**、**T13**、**T17** 和 **T18** 这样的令牌在基线攻击下都被误分类。

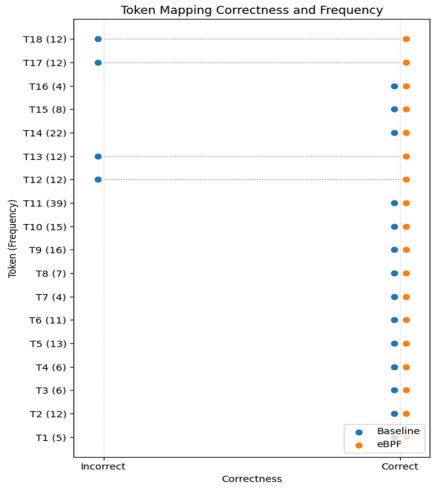

图 1. 有无 eBPF 泄露时各查询的令牌映射准确率

这些令牌的频率均为 12，仅凭结果大小无法将它们彼此区分。

相比之下，增强版 eFMA 方法正确映射了所有令牌，包括基线方法未能映射的令牌。尽管**T12** 和**T13** 的频率相同，但它们在各自的查询中访问了不同的密文文件集合。借助 eBPF，我们能够观测这些文件访问，并将其与明文索引中已知的文档集合进行匹配。这种额外的泄露打破了基线方法无法解决的平局。

像**T11**、**T9**和**T1**这样的其他令牌具有唯一频率，因此两种方法都能正确识别它们，这符合预期。

E. 主要发现

- **文件访问模式泄露的信息比结果大小更多**。只需几行 eBPF 跟踪代码，我们就不仅能恢复返回了多少文档，还能精确确定返回的是哪些文档。
- **系统级泄露显著增强了攻击能力**。恢复率从 77.8% 提升至 100%，表明文件级观测弥补了基于频率的歧义所留下的缺口。
- **动态可搜索对称加密系统即使未经任何修改也容易受到攻击**。我们的增强攻击无需对可搜索对称加密方案进行任何更改。我们将整个系统视为黑盒，并从外部提取所有泄露信息，这表明即使使用经过形式化安全证明的可搜索对称加密方案，实际部署仍可能面临风险。

VII. 未来工作

这项工作还有几个值得探索的方向。首先，尽管我们当前的攻击假设对手能够部分访问明文数据集（以帮助构建索引），但如果能降低这一假设，并探索攻击者在背景知识更有限的情况下能做到什么程度，将会很有价值。

另一个顺理成章的后续方向是在存在更新的情况下运行同样的攻击。动态可搜索对称加密方案通常允许插入和删除文档，研究这些操作会如何影响我们通过 eBPF 观察到的泄露将很有价值。如果更新过程中出现新的模式，它们对攻击者而言也可能同样有用。

我们也希望了解不经意随机存取存储器（ORAM）等机制能否帮助防御此类泄露。ORAM 通过重新排列数据并添加虚假访问来隐藏访问模式，这或许能缓解 eBPF 所依赖的信号。尤其是在攻击者于系统调用级别进行监视时，这种方法是否足够仍是一个悬而未决的问题。

除了文件访问之外，eBPF 还能让我们访问更广泛的系统事件。未来的工作可以研究时序侧信道、内存使用模式或系统调用序列，以发现更多泄露信息。这些信息可以叠加到现有攻击之上，也可以用于增强泄漏滥用类别中的其他攻击。

VIII. 结论

本研究表明，系统级行为可能在可搜索对称加密方案中产生强大的泄漏通道。具体而言，我们展示了在云服务提供商（CSP）一侧使用 eBPF 监视文件访问，会显著增强攻击者恢复查询的能力。

我们利用这些文件级泄露提出的增强攻击，优于传统的基于频率的方法，并在我们的测试数据集上实现了完全的查询恢复。结果凸显了形式化安全模型与现实系统底层实际运行方式之间的差距。

展望未来，SSE 和 DSSE 方案在安全模型中纳入系统级泄露至关重要。eBPF 等工具使攻击者能够轻松观察内核所见的内容，因此不能再忽视这些向量。我们建议未来的工作加强对文件访问泄漏的防御，或可采用 ORAM 等混淆技术；同时，研究人员也应重新审视威胁模型，以反映已部署环境中实际可观测的信息。

参考文献

[1] R. Curtmola、J. Garay、S. Kamara 和 R. Ostrovsky，《可搜索对称加密：改进的定义与高效构造》，载于第*13*届*ACM*计算机与通信安全会议（*CCS*）论文集。ACM，2006年，第79–88页。 [在线]。获取地址：https://doi.org/
10 1145 1180405 1180417 。/。S. Oya 和 F.
[2] Kerschbaum，《隐藏访问模式还不够：利用可搜索加密中的搜索模式泄漏》，载于第*30*届*USENIX*安全研讨会（*USENIX Security 21*）。USENIX协会，2021年，第1463–1480页。 [在线]。获取地址：
https://www.usenix.org/conference/usenixsecurity21/presentation/oya [3] L. Xu、H. Duan、A. Zhou、X. Yuan 和 C. Wang，《解读并缓解可搜索对称加密中的泄漏滥用攻击》，发表于《*IEEE*信息取证与安全汇刊》，第16卷，第5310–5325页，2021年。 [在线]。获取地址：https://doi.org/
10 1109 2021 3128823 。/TIFS。。，J. Jia、Y. Zhu、D.
[4] Williams、A. Arcangeli、C. Canella、H. Franke、T. Feldman-Fitzthum、D. Skarlatos、D. Gruss 和 T. Xu，《使用 eBPF 实现可编程的系统调用安全》，*arXiv*预印本 *arXiv:2302.10366*，2023年。[在线]。获取地址：
https://arxiv.org/abs/2302.10366

[5] B. Gbadamosi、L. Leonardi、T. Pulls、T. Høiland-Jørgensen、S. Ferlin- Reiter、S. Sorce 和 A. Brunstr¨om，《Linux 内核中的 eBPF 运行时》，arXiv预印本 arXiv:2410.00026，2024年。 [在线]。获取地址：https://arxiv.org/abs/ 2410 00026 。D. X. Song、D. Wagner 和 A. Perrig，[6]《加密数据搜索的实用技术》，载于《2000年IEEE安全与隐私研讨会论文集》。S&P 2000。IEEE，2000年，第44–55页。[7] S. Kamara、C. Papamanthou 和 T. Roeder，《动态可搜索对称加密》，载于《2012年ACM计算机与通信安全会议论文集》。ACM，2012年，第965–976页。[8] D. Cash、S. Jarecki、C. S. Jutla、H. Krawczyk、M.-C. Rosu 和 M. Steiner，《超大型数据库中的动态可搜索加密：数据结构与实现》，载于NDSS，2014年。[在线]。获取地址：https://www.ndss-symposium.org/ndss2014/dynamic-searchable-encryption-very-large-databases-data-structures-and-implementation/ [9] M. Islam、M. Kuzu 和 M. Kantarcioglu，《可搜索加密中的访问模式泄露：影响、攻击与缓解》，载于NDSS，2012年。[在线]。获取地址：https://www.ndss-symposium.org/ndss2012/access-pattern-disclosure-on-searchable-encryption-ramification-attack-and-mitigation/ [10] L. Xu、L. Zheng、C. Xu、X. Yuan 和 C. Wang，《针对前向与后向隐私可搜索对称加密的泄漏滥用攻击》，《ACM隐私与安全汇刊》（TOPS），第26卷，第1期，第1–30页，2023年。[11] D. Cash、P. Grubbs、J. Perry 和 T. Ristenpart，《针对可搜索加密的泄漏滥用攻击》，载于《第22届ACM SIGSAC计算机与通信安全会议论文集》。ACM，2015年，第668–679页。[12] G. Gui、S. Hu、Q. Wang、C. Wang 和 K. Ren，《可搜索对称加密中的体积泄露：攻击与缓解措施》，《IEEE信息取证与安全汇刊》，第18卷，第215–230页，2023年。[13] A. Bakas 和 A. Michalas，《Power range：支持范围查询的前向隐私多客户端对称可搜索加密》，载于《第20届安全与密码学国际会议（SECRYPT 2023）论文集》。SciTePress，2023年，第55–66页。[14]《什么是 eBPF？eBPF 技术简介与深入解析》，https://ebpf.io/what-is-ebpf/，无日期，访问日期：2025-05-09。[15] B. Gregg，《BPF性能工具：Linux 系统与应用程序可观测性》，Addison-Wesley Professional，2019年，ISBN：9780136554820。[16] D. Borkmann，《用 BPF 管理网络：Linux 内核邂逅可编程网络》，Linux每周新闻（LWN），2015年。 [在线]。获取地址：https://lwn.net/Articles/740157/ [17] T. Baranski、G. Singh 和 M. Benari，《面向容器的 BPF：它是什么以及如何工作》，博客文章，2021年。 [在线]。获取地址：https://www.redhat.com/en/blog/bpf-containers-what-it-and-how-it-works[18]C. Authors，《Cilium：基于 eBPF 的网络、安全与可观测性》，项目文档，2020年。 [在线]。获取地址：https://cilium.io/blog/2020/06/04/whats-new-in-cilium-18[19] eBPF峰会，《使用 eBPF 实现可观测性与安全》，会议演讲，2023 [ ] 2023 。在线。获[20] 取地址：https://ebpf.io/summit- A. Security，《Tracee：使用 eBPF 实现运行时安全与取证》，GitHub 仓库，2022年。[在线]。获取地址：https://github.com/aquasecurity/tracee [21] H. Lu、S. Wang、Y. Wu、W. He 和 F. Zhang，《Moat：迈向安全的 BPF内核扩展》，载于《第33届USENIX安全研讨会（USENIX Security 24）论文集》。USENIX协会，2024年，第1153–1170页。 [在线]。获取地址：https://www.usenix.org/conference/usenixsecurity24/presentation/lu-hongyi [22] D. K. Nguyen 和 S. Ruj，《可搜索加密中泄漏滥用攻击及其缓解措施的系统综述》，IEEE Access，第11卷，第37 291–37 307页，2023年。[23] DK-Nguyen，《searchable-encryption-database：对称可搜索加密（SSE）数据库的一种实现》，https://github.com/DK-Nguyen/searchable-encryption-database，无日期，访问日期：2025-05-09。[24] E. Corp 和 W. W. Cohen，《安然邮件数据集》，https://www.loc.gov/item/2018487913/，2015年，美国联邦能源监管委员会编制。[25] A. Robertson 和 bpftrace 贡献者，《bpftrace：面向 Linux 的高级跟踪语言》，https://github.com/bpftrace/bpftrace，无日期，访问日期：2025-05-09。