New Constructions for Forward and Backward Private Symmetric Searchable Encryption 阅读笔记

李忆诺

2024 年 4 月 2 日

1 基本信息

1.1 论文来源

Chamani J G , Papadopoulos D , Papamanthou C ,et al.New Constructions for Forward and Backward Private Symmetric Searchable Encryption[C]//the 2018 ACM SIGSAC Conference.ACM, 2018.DOI:10.1145/3243734.3243833.

1.2 概述

本文基于动态对称可搜索加密问题，在 Bost 等人工作的基础上，引入了三种新的结构，以多种方式改进了以前的结果。第一种方案实现了 type-ii 型后向隐私；第二种方案实现了最强的后向隐私级别 (type-i)；最终方案在第二种方案的基础上进行了改进，以额外的泄漏为代价减少了搜索的往返次数 (type-iii)。

2 论文要点

2.1 背景

当前的动态对称可搜索加密方案中，对于限制后向隐私泄露的方案在文献中研究得少得多。Bost 等人引入了后向隐私的正式定义，其中从最安全到最不安全定义了三种不同类型的泄露，并提出了四种后向私有方案。本文则在该方案的基础上进行优化改进。

2.2 价值

本文提出了三种具有前向和后向隐私的 SSE 方案，在几个方面改进了 Bost 等人的结果：

- Mitra：快速 type-ii 型，它达到了与 Fides 相同的性能，且其搜索计算时间为 145-253x，更新计算时间为 86-232x，成为现有最高效的前向和后向私有 SSE；
- 最优搜索时间：最优搜索时间 $O(n_w)$ 或准最优搜索时间 $O(n_w \cdot ploy\log(N))$，并给出了两个具有准最优搜索时间的方案 Orion 和 Horus

- Orion：最强后向隐私级别，type-i，它需要 $O(\log N)$ 轮交互，搜索过程需要 $O(n_w \log^2 n)$ 步。
- Horus：type-iii 型，该方案提高了 Orion 的效率，实现了更好的搜索性能并减少了交互轮数。

2.3 问题陈述

本文在 Bost 的工作的基础上，提出了对前向后向私有方案的优化策略，通过引入 $ORAM$ 方案，设计了 Mitra；然后定义了最优搜索时间和非平凡交互，由此提出方案 Orion 和 Horus。

最优搜索时间：如果一个方案的搜索时间复杂度为 $O(n_w)$（或 $O(n_w \cdot poly \log(N))$），我们称该动态对称可搜索加密方案具有最优搜索时间（准最优搜索时间）。

非平凡交互：如果一个方案在搜索过程中需要 $O(n_w)$ 次交互，我们说该方案满足非平凡交互；否则我们说该方案满足平凡交互。

2.4 方法

2.4.1 Mitra

更新

本地维护一个表，用于存储每个关键字下更新条目的数量。对于每一个更新操作，需要进行如下步骤：

- 客户端使用伪随机函数 G，根据关键字 w 和当前更新操作对应次数 $FileCnt[w]$ 计算出一个随机数 $addr$，作为存储到服务器上的标识符；
- 客户端使用同样的伪随机函数 G，用类似步骤计算出随机数，并将其与 $(id||op)$ 进行异或，作为 val，存储至服务器上；
- 由于对其中的 $FileCut[w]$ 进行了处理，因此服务器无法计算得到 $(id||op)$，也无法判断当前条目属于哪个关键字 w。

搜索

对于每一次搜索操作，需要进行如下步骤：

- 客户端计算出 w 的所有对应 $addr$，发给服务器；
- 服务器返回所有 $addr$ 对应的 val；
- 客户端通过异或运算得到 $(id||op)$，最终剥离出 id。

2.4.2 Orion

obvious map

通过使用 $OMAP$ 结构来对 $AVL\ \ Tree$ 的各个节点进行随机化。此时客户端只需存储 $OMAP$ 的 $rootID$，缩小了存储空间。

更新

本地维护一个表，用于存储每个关键字下文件的数量，并存储每个关键词最新一次的标识符 $updt_{cnt}$。更新操作和查询操作分别存储在两个 $obvious \quad map$ 结构中，其 $rootID$ 存储在本地，服务器存储两个 $OMAP$。其中，在 $OMAP_{upd}$ 中存储从 (w, id) 到 $updt_{cnt}$ 的映射，而在 $OMAP_{src}$ 中存储从 $(w, updt_{cnt})$ 到 id 的映射。

其关键在于删除操作：将当前关键字的最新项与待删除项进行交换。这保证了在任何给定时间内，都可以通过搜索 $(w,1),(w,2),...(w,n_w)$ 来检索包含 w 的当前文档对应标识符。

搜索时只需根据本地 $updt_{cnt}$ 依次遍历查找即可。在访问结束后，需要重新映射、重新加密整个访问的子树，并将其存储在 $ORAM$ 中。在查询操作时，可插入一系列虚拟 $OMAP$ 查询操作，以混淆真正携带数据的查询条目。

2.4.3 Horus

由于 Orion 方案即使没有删除操作，其搜索成本也保持不变，即实际上该方案的搜索时间为 $\Theta(n_w \cdot \log^2 n)$，因此本文在 Orion 方案上进行改进，提出了 Horus。

其修改部分如下：

- 使用 $Path-ORAM$ 结构代替 $OMAP_{src}$；
- 为了避免在客户端存储大小为 N 的位置映射，此处使用安全 PRF 生成，并且引入一个计数器 acc_{cnt} 来避免 PRF 的确定性引发的泄露。
- 将 $updt_{cnt}$ 和 acc_{cnt} 一起存放在 $OMAP_{upd}$ 中。

在搜索过程中，客户端每次通过猜测 acc_{cnt} 的值，来查询每个文件，通过二分查找来获得该关键字的最大 acc_{cnt} 值。

2.5 结果

2.5.1 Mitra

Mitra 的更新操作复杂度渐进为 $O(1)$，搜索操作渐进为 $O(a_w)$，需要一次往返来检索 w 的文件标识符，检索实际文件则需要再进行一轮交互。

在 Mitra 方案中，删除操作并没有实现真正意义上的删除。可以执行周期性的“清理”操作，通过让客户端在搜索后删除已删除的条目，重新加密剩余的条目并将其发送回服务器来实现。但 Mitra 的加密方案是确定性的，这将产生相同的密文，泄露信息。为了避免这种情况，可以维护一个额外的计数器映射 $SrcCnt$，它在每次搜索后递增，并作为伪随机函数的输入。

2.5.2 Orion

Orion 的更新复杂度为 $O(\log^2 N)$，其采用的 $obvious \quad map \quad Construction$ 可以处理“批处理”查询，批量执行 n_w 次查询需要 $O(n_w \cdot \log^2 n)$ 时间，满足准最优搜索时间。并且该方案需要 $O(\log N)$ 轮交互，满足非平凡交互。

除此之外，也可以将 $updt_{cnt}$ 和 $Last_{ind}$ 的存储外包给服务器，并在每个操作中下载它，这样，Orion 的本地存储减小至 $O(1)$。而对于服务器来说，由于其存储大小为初始设定的固定值，且存储介质的相对低成本和高可用性，作者不认为这是该方案的严重限制因素。

由于 Orion 使用了黑盒 $obvious \quad map$，且进行了假查询填充，每次更新后访问的 $ORAM$ 条目都被重新映射到新的位置，因此，其满足前向安全，且具有最小的泄露，后向隐私类型为 type-i。

2.5.3 Horus

Horus 由于添加了一个 acc_{cnt} 用于消除 PRF 的确定性带来的泄露问题，因此需要计算 w 对应的最大 acc_{cnt} 值，该进程在 $O(\log d_w)$ 轮后终止，并需要 $O(n_w \cdot \log d_w)$ 次 $ORAM$ 访问。

因此，Horus 的搜索时间和通信复杂度为 $O(n_w \cdot \log d_w \cdot \log N)$，且在搜索期间访问的一些 $ORAM$ 位置与之前在更新期间访问的相同，这引入了一些泄露，使得 Horus 满足后向隐私的 type-iii 类型。

3 评论

3.1 局限性

本文通过 $ORAM$ 方案开发出具有准线性搜索时间和非平凡交互的方案，但 $ORAM$ 方案仅适用于前向私有方案。是否可以提出一种不依赖于 $ORAM$ 的方案，使其具有准线性搜索时间和非平凡交互。

除此之外，对于支持删除的结构来说，如何设计一个具有准最优搜索时间的非平凡交互方案，也是一个悬而未决的问题。

3.2 扩展阅读

Dynamic Searchable Symmetric Encryption with Forward and Stronger Security

Dynamic Searchable Symmetric Encryption With Strong Security and Robustness

3.3 启示

本文的方案为隐私保护提供了更高效和更灵活的解决方案。特别是第一个方案，它实现了 Type-ii 后向隐私，搜索计算时间比以前的构造快了 145 – 253 倍。这对于在实际应用中处理大规模数据集时非常有用。

除此之外，本文中的方案提出了关于**最优搜索时间**和**非平凡交互**两个定义，为后续设计后向私有方案的效率分析制定了一个可参考的标准，将该效率定义与后向隐私泄露程度定义结合起来，为后向安全方案设计提供了理论依据和指导。