结果模式隐藏的联合查询加密多映射方案

杜若飞

（华南师范大学 计算机学院，广东 广州 510631）

摘　要：联合查询加密多映射技术（cEMM）在保障外包数据机密性的同时支持多标签联合查询。为了解决查询结果模式泄漏带来的数据隐私风险，设计了一种具备结果模式隐藏特性的联合查询加密多映射方案。通过优化查询令牌与服务器存储结构，可在不泄漏查询结果模式的条件下完成联合查询。安全性分析表明，该方案能有效隐藏查询结果模式，通信开销相较于较现有代表性方案降低了69.75%。

关键词：对称可搜索加密；加密多映射；联合查询；访问模式；通信效率

DOI：10. 11907/rjdk. 261021　　扫描二维码阅读全文：

中图分类号：TP3　　**文献标识码：**A　　**文章编号：**1672-7800（XXXX）0XX-0001-07

Encrypted Multi-Map Scheme with Result Pattern Hiding for Conjunctive Quevies

DU Ruofei

（*School of Computer Science*，*South China Normal University*，*Guangzhou* 510631，*China*）

Abstract： Joint Query Encryption Multi Mapping Technology （cEMM） supports multi label joint queries while ensuring the confidentiality of outsourced data. In order to address the data privacy risk caused by the leakage of query result patterns，a joint query encryption multi mapping scheme with result pattern hiding characteristics was designed. By optimizing the query token and server storage structure，joint queries can be completed without leaking the query result pattern. Security analysis shows that this scheme can effectively hide the query result pattern，and the communication overhead is reduced by 69.75% compared to existing representative schemes.

Key Words： symmetric searchable encryption；encrypted multi-map；conjunctive search；access pattern；communication efficiency

0　引言

对称可搜索加密（Searchable Symmetric Encryption，SSE）使用户能安全地将加密数据外包至云服务器，并在不暴露明文的前提下对密文执行查询。Song等[1]提出SSE方案以来，该领域的研究已持续进行了二十多年[2-12]。早期，研究集中于单关键字查询场景，虽在查询效率上优势显著，但表达能力有限，难以应对日益复杂的实际检索需求[3-9]。Cash等[13]提出OXT方案，在次线性时间内高效支持联合查询，成为了后续研究的重要基础。

目前，针对SSE泄露信息的攻击手段不断涌现，这些信息可能在实践中被敌手加以利用，从而推断出用户的查询内容[14,15]。该问题在更复杂的联合查询场景中更加严峻，例如关键字对结果模式（Keyword Pair Result Pattern，KPRP）、跨查询交集结果模式（Intersection Result Pattern，IP）等泄露，被认为会显著削弱系统隐私性[16,17]。因此，如何抑制信息泄露，便成为了联合查询可搜索加密设计中亟需解决的核心问题。

目前，研究者探索了多种泄露抑制技术。Lai等[16]引入一种新的密码学原语——对称隐藏向量加密（Symmetric Hidden Vector Encryption，SHVE），并在此基础上提出HXT方案。HXT有效避免了KPRP泄露，但需额外一轮交互且无法消除IP泄露。Wang等[17]提出基于异或过滤器（Xor Filter）的Doris方案，在非交互场景下同时隐藏了KPRP与IP。然而，上述方案普遍存在通信开销较高的问题，在大规模应用场景中面临一定的效率瓶颈。因此，如何在有效隐藏访问模式的同时降低通信开销、提升查询效率，仍是一个开放性问题。

为此，本文提出一种结果模式隐藏的联合查询加密多映射方案（Conjunctive Encrypted Multi-Map，cEMM），在保持强泄露抑制能力的同时有效降低通信与计算开销。主要贡献为：①针对联合查询中的泄露问题，形式化定义了更强的结果模式隐藏方案，并证明其蕴含KPRP隐藏；②设

收稿日期：2026-01-15

作者简介：杜若飞（1999-），男，CCF会员，华南师范大学计算机学院硕士研究生，研究方向为可搜索加密。

计了高效、安全的RH-cEMM方案，基于令牌派生机制有效避免了现有方案中搜索令牌数量的乘性膨胀，将标签—值包含关系抽象为布尔值并进行随机掩码异或加密，可有效隐藏查询结果模式。

对于联合查询$\varphi = l_1 \wedge ... \wedge l_m$，本文所提方案与现有代表性方案的理论性能如表1所示。其中，m为查询标签数量；n_{l_1}为l_1对应值列表的大小；n_φ为整个查询结果集合的大小；N为任意两个标签交集集合大小之和，即$N = \sum_{l_i, l_j \in \mathbb{L}} |MM[l_i] \wedge MM[l_j]|$；$n$表示所有标签相应值集合大小之和，即$n = \sum_{l_i \in \mathbb{L}} |MM[l_i]|$；$L$表示标签总数。

Table 1 Theoretical performance of the four schemes

表1 4种方案的理论性能

方案	通信复杂度	搜索时间复杂度	服务器存储	KPRP隐藏	IP隐藏	结果模式隐藏	通信轮数
STE-BQ[18]	$\mathcal{O}(n_\varphi)$	$\mathcal{O}(n_{l_1})$	$\mathcal{O}(L+n)$	√	×	×	1
vRHS[19]	$\mathcal{O}(n_{l_1} \cdot m + n_{l_1})$	$\mathcal{O}(n_{l_1} \cdot m)$	$\mathcal{O}(n)$	√	√	√	2
Doris[17]	$\mathcal{O}(n_{l_1} \cdot m + n_\varphi)$	$\mathcal{O}(n_{l_1} \cdot m)$	$\mathcal{O}(N)$	√	√	×	1
RH-cEMM	$\mathcal{O}(m + n_{l_1})$	$\mathcal{O}(n_{l_1} \cdot m)$	$\mathcal{O}(N)$	√	√	√	1

1 相关工作

可搜索加密(Searchable Encryption, SE)的早期研究可追溯至Song等人提出的方案[1]。该方案在密文数据上执行关键字搜索，但搜索开销与数据库的文档数量呈线性相关，难以满足大规模数据场景下的效率需求。Curtmola等[20]提出基于倒排索引结构的SSE方案，系统给出了适应性安全(Adaptive Security, AS)的形式化定义，为后续SSE方案的安全分析奠定了理论基础。Goh[21]提出基于Bloom过滤器的Z-IDX索引结构，相较于传统倒排索引，在索引存储方面更紧凑。

为了提升可搜索加密方案的功能性，尤其是对多关键字联合查询的支持能力，Cash等提出OXT方案[13]。通过构造双索引结构(TSet与XSet)，结合基于Diffie-Hellman的安全计算操作，实现了次线性时间复杂度下的多关键字交集查询。然而，OXT存在严重的泄露问题，包括KPRP、IP。因此，后续大量研究围绕如何避免泄露、扩展查询功能、降低计算与存储开销等方面展开。

Kamara等[22]提出IEX方案，在初始化阶段预先存储任意两个关键字对应的文档标识符交集，从而在查询阶段直接返回结果，显著提升了联合查询效率。Lai等[16]提出HXT方案，通过结合SHVE有效消除KPRP泄露，使敌手无法通过结果模式推断关键字对之间的关联关系，但需额外的一轮交互，且仍无法避免IP泄露。Patel等[23]提出的ConjFilter方案，成功消除了OXT单关键字查询中的搜索模式泄露及多关键字交集查询中的IP泄露问题。相较于IEX，ConjFilter同样依赖预存储关键字对之间的交集，但通过精巧的索引构造与令牌设计，避免了关键字泄露，在安全性上优于OXT与IEX。

鉴于联合查询的本质可归结为集合交集计算，部分研究工作尝试引入隐私集合求交(Private Set Intersection, PSI)技术，构造同时满足KPRP与IP隐藏的联合查询方案[19]。此类方法在安全性方面表现突出，但依赖复杂的密码学协议，导致客户端在计算与存储方面负担较重。Wang等[17]提出Doris方案，基于Xor Filter构建索引结构，在非交互式查询场景下实现了对KPRP与IP的同时隐藏。然而，该方案在通信开销方面仍然较高，限制了其在实际系统中的进一步应用。

2 预备知识

2.1 IND-CPA安全对称加密方案

定义1 设对称加密方案SE由($KeyGen$, Enc, Dec)等3个算法组成[23]。其中，密钥生成算法$KeyGen$以安全参数λ为输入，生成密钥$k \in \boldsymbol{K}$。加密算法Enc以密钥k、明文$m \in \{0,1\}^*$为输入，输出对m加密后的密文c。若对于所有PPT敌手$\mathcal{A}$，可忽略以下优势，则称SE满足IND-CPA安全。

$$\mathbf{Adv}_{\Sigma,A}^{IND-CPA}(\lambda) = \Pr[\mathcal{A}^{\mathcal{O}(K,0,\cdot,\cdot)}(1^\lambda) = 1] - \Pr[\mathcal{A}^{\mathcal{O}(K,1,\cdot,\cdot)}(1^k) = 1] \tag{1}$$

2.2 联合查询加密多映射

多映射(multi-map, MM)数据结构支持存储标签/值对$MM = \{(l, \vec{v})\}$，l为标签且$l \in \mathbb{L}$；$\vec{v}$为标签l对应的值列表且$\vec{v} \in \boldsymbol{V}$。令$\vec{v}[i]$表示$\vec{\boldsymbol{v}}$中第$i$个值，$MM[l]$表示标签$l$对应的值列表，多映射$MM$支持以下操作：①$Get(l_i)$操作输入标签$l_i$，输出标签$l$对应的值列表$\vec{v}_i = MM[l]$；②$Put(l_i, \vec{v}_i)$操作输入标签$l_i$与值列表$\vec{v}_i$，并将$(l_i, \vec{v}_i)$插入多映射。

为了扩展多映射的实用性，还考虑了支持联合查询的多映射(Conjunctive MM, cMM)，即对一个多标签联合查询$\varphi = l_1 \wedge ... \wedge l_m$，返回所有标签对应值列表的交集$MM[\varphi] = MM[l_1] \wedge ... \wedge MM[l_m]$。接下来，给出联合查询加密多映射的形式化定义。

定义2 设多映射$MM = \{(l, \vec{v})\}$，将文献[17]的联合查询加密多映射方案形式化定义为：①$Setup(1^\lambda, MM)$算法由客户端运行，输入安全参数λ、多映射MM，输出加密多映射$cEMM$、对应密钥K；②$Search(K, \varphi, cEMM)$算法由客户端与服务器交互运行，客户端输入联合查询$\varphi = l_1 \wedge ... \wedge l_m$、密钥$K$，服务器输入加密多映射$cEMM$，最

终客户端得到查询结果 $MM[\varphi]$。

定义 3　本文采用联合查询加密多映射的安全性定义，令 $\Sigma=(Setup, Search)$ 为联合查询加密多映射，$\mathcal{L}$ 为泄露函数[17]。同时，定义以下两个实验：

（1）$Real^{\Sigma}_{\mathcal{A}}$。首先，敌手 $\mathcal{A}$ 选择一个多映射 MM 交给实验，实验运行 $Setup(1^{\lambda}, MM)$ 并将生成的 $cEMM$ 返回给 $\mathcal{A}$；其次，$\mathcal{A}$ 自适应地选取查询 φ 交给实验，实验运行 $Search(K, \varphi, cEMM)$ 并将搜索过程的状态与结果返回给 $\mathcal{A}$；最后，$\mathcal{A}$ 输出比特 b 作为实验输出。

（2）$Ideal^{\Sigma}_{\mathcal{A},\mathcal{S}}$。首先，敌手 $\mathcal{A}$ 选择一个多映射 MM 交给实验，实验运行 $\mathcal{S}(\mathcal{L}(MM))$ 并返回 $cEMM$ 给 $\mathcal{A}$；其次，$\mathcal{A}$ 自适应地选取查询 φ 交给实验，实验运行 $S(\mathcal{L}(MM, \varphi))$ 并将搜索过程的状态与结果返回给 $\mathcal{A}$；最后，$\mathcal{A}$ 输出比特 b 作为实验输出。

若对于任意多项式时间敌手 $\mathcal{A}$，均有一个仿真器 $\mathcal{S}$ 存在以下关系，则称 Σ 在泄露函数 $\mathcal{L}$ 下是自适应安全的。

$$\Pr[\mathbf{Real}^{\Sigma}_{\mathcal{A}}(\lambda)=1]-\Pr[\mathbf{Ideal}^{\Sigma}_{\mathcal{A},\mathcal{S}}(\lambda)=1]\leqslant negl(\lambda) \quad (2)$$

2.3　TSet

TSet 是由 Cash 提出的一种加密多映射方案，由以下 3 个多项式时间算法组成[13]。

（1）$TSet.Setup(1^{\lambda}, T)$ 算法。以安全参数 λ 与多映射 $T=\{(l,(v_1,...,v_{T[l]}))\}_{l\in\mathbb{L}}$ 为输入，$\mathbb{L}$ 为 T 中所有标签 l 的集合，v_i 为标签 l 对应的所有值，最终输出为密钥 K_T、加密多映射 TSet。

（2）$TSet.GetTag(K_T, l)$ 算法以密钥 K_T 与查询标签 l 为输入，输出结果为查询令牌 $stag$。

（3）$TSet.Retrieve(TSet, stag)$ 算法输入查询令牌 $stag$ 与加密多映射 TSet，输出查询结果，即标签 l 对应的值列表 $T[l]$。

2.4　结果模式隐藏

定义 4　给定 $cEMM$ 方案 Σ，定义游戏 $Game^{\mathcal{L}}_{\mathcal{A}}(n,b)$ 为：

（1）敌手 $\mathcal{A}$ 选择 MM_0、MM_1 交给挑战者 $\mathcal{C}$，这两个多映射满足标签集 $\mathbb{L}$ 大小相等，同为 n。任意标签 $l\in\mathbb{L}$，$|MM_0[l]|=|MM_1[l]|$。挑战者 $\mathcal{C}$ 均匀抽样比特 b，运行算法 $\Sigma.Setup(\mathcal{L}(MM_b))$，并将生成的 $cEMM$ 发送给敌手 $\mathcal{A}$。

（2）敌手 $\mathcal{A}$ 自适应地选取查询 φ，对每个查询挑战者 $\mathcal{C}$ 运行算法 $\Sigma.Search(\mathcal{S}(\mathcal{L}(MM,\varphi)))$，并将输出结果返回至敌手 $\mathcal{A}$。最终，$\mathcal{A}$ 输出比特 b'。

定义 5　令 Σ 为 $\mathcal{L}$-自适应安全的联合查询加密多映射方案，如果对于任意多项式时间敌手 $\mathcal{A}$，满足以下等式，则称 Σ 满足结果模式隐藏。

$$\Pr[Game^{\mathcal{L}}_{\mathcal{A}}(n,0)=1]=\Pr[Game^{\mathcal{L}}_{\mathcal{A}}(n,1)=1] \quad (3)$$

3　结果模式隐藏的加密多映射方案

3.1　Doris 方案分析与改进

联合查询的目标为：在加密多映射中检索同时属于多个标签 $l_1,...,l_m$ 的值。首先，由服务器检索第一个标签 l_1 的候选值集合 $MM[l_1]$；其次，验证每个候选值是否同时属于其他标签，以过滤不满足条件的值。然而，若服务器对候选值逐标签验证，会泄露值与各标签之间的包含关系，即 KPRP。为此，Doris 延续了 HXT 方案的“整体验证”机制，使服务器只能获得每个值是否同时满足所有查询标签的最终布尔结果，而无法获知其与单个标签的包含关系。该机制有效隐藏了 KPRP，但代价是客户端需为每个候选值生成针对其余标签的过滤令牌，会使令牌数量与 $|MM[l_1]|\cdot m$ 呈线性增长，在多标签或高频标签场景下引发通信开销的显著膨胀。此外，“整体验证”机制无法隐藏最终的查询结果模式，服务器仍可直接观察最终哪些候选值满足“整体”查询结果，仍存在潜在的隐私泄露风险。

为此，RH-cEMM 方案从令牌结构与验证粒度层面进行结构性改进。在令牌层面，设计了一种高效的令牌派生机制，客户端仅针对每个查询标签发送单个令牌 $xtoken$，服务器利用本地预存的伪随机值 y，即可自主派生出所有值对应的过滤令牌 $xtag$，使令牌数量仅与 m 相关，有效降低了令牌开销。在验证层面，不再采用“整体验证”思路，而将每个值与各标签的包含关系抽象为布尔值 $b=\{0,1\}$，$b=1$ 表示标签包含值，反之不包含，并通过随机掩码对其进行异或加密，由客户端进行本地验证。该方法不仅有效隐藏了单个值与标签的关联，还能隐蔽查询结果模式。

3.2　RH-cEMM 方案设计

RH-cEMM 方案包括初始化、服务器端搜索两个核心算法。算法 1 为初始化算法 $Setup$，算法输入安全参数 λ 与多映射 $MM=\{(l,\vec{v})\}_{l\in\mathbb{L}}$，输出用户私钥 K 与加密多映射 $cEMM$（由 $TSet$、$XMap$ 两部分组成）。首先，对于给定标签 l，得到多映射 MM 中与标签 l 关联的每个值 $\vec{v}[i]$，计算加密值 $e\leftarrow SE.Enc(k_l,\vec{v}[i])$ 与 $y\leftarrow F(K_1, l\|i)$，并将 (e,y) 加入集合 t；其次，针对任意不同于 l 的标签 l'，若值 $\vec{v}[i]$ 同时属于 $MM[l']$，则在 $XMap$ 中设置 $XMap[\text{xtag}]\leftarrow 1\oplus \text{F}(K_2, l\|l'\|i)$，否则设置 $XMap[\text{xtag}]\leftarrow 0\oplus F(K_2, l\|l'\|i)$，以不可被区分的方式编码不同标签之间的交集关系；最后，输出主密钥集合 $K=(K_E, K_X, K_T)$、加密多映射 $cEMM=(TSet, XMap)$。

算法 1　初始化算法 $Setup$

输入：λ、$MM=\{(l,\vec{v})\}_{l\in\mathbb{L}}$。

输出：K、$cEMM$。

1.$K_E, K_X, K_1, K_2\leftarrow\{0,1\}^{\lambda}$

2.$XMap\leftarrow\varnothing$

3.FOR $l\in\mathbb{L}$

4. $t\leftarrow\varnothing$, $k_l\leftarrow F(k_E, l)$

5. FOR $i=1$ *to* $|MM[l]|$

6. $e\leftarrow SE.Enc(k_l,\vec{v}[i])$, $y\leftarrow F_1(K_1, l\|i)$, $t\leftarrow t\cup(e,y)$

7. FOR $l'\in\mathbb{L}\backslash l$

8. $xtag\leftarrow F_1(K_X, l\|l')\oplus y$

9. IF $\vec{v}_l \in MM[l']$ THEN
10. $XMap[xtag] \leftarrow 1 \oplus F_2(K_2, l \| l' \| i)$
11. ELSE
12. $XMap[xtag] \leftarrow 0 \oplus F_2(K_2, l \| l' \| i)$
13. END IF
14. END FOR
15. END FOR
16. $T[l] \leftarrow t$
17.END FOR
18.$(TSet, K_T) \leftarrow TSet.Setup(1^\lambda, T)$
19.RETURN $\{K = (K_E, K_X, K_T), cEMM = (TSet, XMap)\}$

算法2为搜索算法*Search*。在查询阶段，客户端先对标签l_1生成*TSet*查询令牌*stag*，并针对每个标签$l_2 ... l_m$计算过滤令牌$xtoken_1, ..., xtoken_m$。服务器接收到令牌后，调用$t \leftarrow TSet.Search(Stag)$得到与标签$l_1$对应的元组列表$t$，并结合$t$中预存储的$y$，为每个密文候选值$e$派生对应的过滤令牌$xtag_1, ..., xtag_m$。随后，利用令牌检索*XMap*，以取回对应的加密布尔值$r[j]$，为每个密文值e生成一个长度为$m-1$加密布尔值列表r。由于所有布尔值都被随机掩码进行了异或加密，因此服务器无法获知其明文内容。最后，客户端在本地解掩码所有加密布尔值$b \leftarrow r[j] \oplus F(K_X, l_1 \| l_j \| i)$，根据$b_i$值过滤对应的密文值$e_i$，得到满足联合查询条件的最终结果集合。

算法2 搜索算法*Search*

输入：K、$\varphi = l_1 \wedge ... \wedge l_m$、*cEMM*。
输出：R。
Client：
1.$stag \leftarrow TSetGetTag(l_1, K_T)$
2.FOR $i = 1$ *to* $m-1$
3. $xtoken_i \leftarrow F_1(K_X, l_1 \| l_{i+1})$
4.END FOR
5.send $stag, xtoken_1, ..., xtoken_m$ to server
Server：
6.$t \leftarrow TSet.Search(Stag)$
7.$res \leftarrow [\,]$
8.FOR $i = 1$ *to* $|t|$
9. $r \leftarrow [\,]$, $(e, y) \leftarrow t[i]$
10. FOR $j = 1$ *to* $m-1$
11. $xtag \leftarrow y \oplus xtoken_i$
12 $r[j] \leftarrow XMap[xtag]$
13. END FOR
14. $res[i] \leftarrow (e, r)$
15.END FOR
16.send *res* to Client
Client：
17.FOR $i = 1$ *to* $|res|$
18. $(e, r) \leftarrow res[i]$
19. $flag \leftarrow true$
20. FOR $j = 1$ *to* $m-1$
21. $b \leftarrow r[j] \oplus F_2(K_2, l_1 \| l_{j+1} \| i)$
22. IF $b = 0$ THEN
23. $flag \leftarrow false$
24. BREAK
25. END IF
26. IF $flag = true$ THEN
27. $R \leftarrow R \cup SE.Dec(K_l, e)$
28.END FOR
29.RETURN

3.3 安全性分析

本文假设$Q = \{\varphi_1, ..., \varphi_m\}$表示一系列联合查询请求，每个查询都由$m$个标签组成，即$\varphi_i = l_1 \wedge ... \wedge l_m$。对于查询$\varphi_i$与其首个标签$\varphi_i[1] = l_1$，令$SP(\varphi_i[1])$表示$\varphi_i[1]$的重复模式；令$RP(\varphi_i)$表示第一个标签的查询结果模式；令$N$表示每个标签值列表的长度。

$$SP(\varphi_i[1]) = \{j \mid \varphi_j[1] = \varphi_i[1] \varphi_j \in Q\} \tag{4}$$

$$RP(\varphi_i) = MM[\varphi_i[1]] \tag{5}$$

$$N(\varphi_i[j]) = |MM[\varphi_i[j]]| \tag{6}$$

定理1 设泄露函数$\mathcal{L}_{cEMM}(MM, Q)$以多映射*MM*与查询序列$Q$为输入参数，输出$(N, SP, RP)$。设$F$为安全的伪随机函数，*SE*为IND-CPA安全的对称加密方案，TSet实例是自适应安全的，那么加密多映射方案RH-cEMM具有$\mathcal{L}_{cEMM}$—自适应安全[24]。

证明： 采用混合论证形式构建一系列中间游戏$G_0, ..., G_n$、仿真器$\mathcal{S}$，其中G_0的分布与$Real_{\mathcal{A}}^{\Sigma}$分布相同，相邻游戏仅作微小修改，确保相邻游戏在计算上不可区分。在最终的游戏中，仿真器仅利用泄露函数中的信息，构造与真实实验不可区分的理想实验$Ideal_{\mathcal{A},\mathcal{S}}^{\Sigma}$，从而完成证明。

（1）构造游戏G_1。G_0中e的生成方法修改为$e \leftarrow SE.Enc(k_l, 0^\lambda)$。由于*SE*满足IND-CPA安全性的对称加密方案，因此存在一个多项式时间敌手$\mathcal{B}_1$使得：

$$|\Pr[G_1 = 1] - \Pr[G_0 = 1]| \leqslant negl(\lambda) \cdot Adv_{SE,\mathcal{B}_1}^{IND-CPA}(\lambda) \tag{7}$$

（2）构造游戏G_2。修改G_1中与TSet相关的代码，即将$stag \leftarrow TSetGetTag(l_1, K_T)$与$(TSet, K_T) \leftarrow TSet.Setup(1^\lambda, T)$替换为通过TSet仿真器$S_T$生成的$(TSet, Stags) \leftarrow \mathcal{S}_T(\mathcal{L}(T, SP, RP, N))$。由于TSet实例在自适应攻击下满足语义安全，因此存在敌手$\mathcal{B}_2$使得：

$$|\Pr[G_2 = 1] - \Pr[G_1 = 1]| \leqslant Adv_{\mathcal{B}_2}^{\mathrm{TSet}}(\lambda) \tag{8}$$

（3）构造游戏G_3。G_3与G_2基本一致，主要有两处改变：①G_3通过维护映射Y计算y，若$Y[l,i]$存在，直接返回其值，否则从$\{0,1\}^\lambda$中均匀抽样，将新值存入$Y[l,i]$并返回；②G_3通过维护映射X计算xtag，若$X[l,l',i]$存在，直接返回其值，否则从$\{0,1\}^\lambda$中均匀抽样，将新值存入$X[l,l',i]$并返回。

同时，不再执行$XMap[xtag] \leftarrow 0 \oplus F(K_2, l \| l' \| i)$，而改为均匀选取一个比特$b$，设置$XMap[xtag] \leftarrow b$。由于$F$为安全的伪随机函数，因此存在敌手$\mathcal{B}_3$使得[25]：

$$|\Pr[G_3=1]-\Pr[G_2=1]|\leqslant 3\cdot Adv^{PRF}_{\mathcal{B}_1,F}(\lambda) \quad (9)$$

(4)构造游戏 G_4。G_4 与 G_3 的区别在于 *xtoken* 的构建方式,G_4 不再计算 $xtoken_i \leftarrow F(K_X, l_1\|l_{i+1})$,而直接利用 X 与 Y 计算 *xtoken*,即在计算 *xtoken* 时直接返回值 $Y[l,i]\oplus X[l,l',i]$。由于仅改变构建方式,因此存在敌手 $\mathcal{B}_4$ 使得:

$$|\Pr[G_4=1]=\Pr[G_3=1]| \quad (10)$$

综上,$Ideal^{\Sigma}_{\mathcal{A},\mathcal{S}}$ 的构建过程为:在仿真实验中,仿真器 $\mathcal{S}$ 仅接收泄露函数 $\mathcal{L}_{cEMM}(MM,Q)$,并输出相应的仿真视图,如算法3所示。此时仿真视图与 G_4 无法区分,因此存在:

$$|\Pr[G_4=1]|=\Pr[\mathrm{Ideal}^{\Sigma}_{\mathcal{A},\mathcal{S}}(\lambda)=1]| \quad (11)$$

关联以上游戏可得:

$$|\Pr[\mathrm{Real}^{\Sigma}_{\mathcal{A}}(\lambda)=1]-\Pr[\mathrm{Ideal}^{\Sigma}_{\mathcal{A},\mathcal{S}}(\lambda)=1]|\leqslant negl(\lambda) \quad (12)$$

算法3 仿真器 $\mathcal{S}$

输入:$\mathcal{L}_{cEMM}(MM,Q)$。

输出:仿真视图。

1.Setup(MM,Q):

2.FOR $l\in\varphi[1]$ where $\varphi\in Q$

3. FOR $i=1$ *to* $|MM[l]|$

4. $e\leftarrow SE.Enc(k_l,0)$, $y\leftarrow\{0,1\}^{\lambda}$

5. $Y[l,i]\leftarrow y$, $t\leftarrow t\cup(e,y)$

6. FOR $l'\in\mathbb{L}\backslash l$

7. IF $X[l,l',i]=\perp$ THEN

8. $X[l,l',i]\leftarrow\{0,1\}^{\lambda}$

9. END IF

10. $xtag\leftarrow X[l,l',i]$

11. $XMap[xtag]\leftarrow\{0,1\}$

12. END FOR

13. END FOR

14. $T[l]\leftarrow t$

15.END FOR

16.$(TSet,Stags)\leftarrow\mathcal{S}_T(\mathcal{L}(T,SP,RP,N))$

17.Search(Q):

18.FOR $i=1$ *to* $m-1$

19. $xtoken_i\leftarrow Y[l,i]\oplus X[l,l',i]$

20.END FOR

21.$t\leftarrow\mathcal{S}_T(\mathcal{L}(stag,TSet))$

22.$res\leftarrow[\,]$

23.FOR $i=1$ *to* $|t|$

24. $r\leftarrow[\,]$

25. $(e,y)\leftarrow t[i]$

26. FOR $j=1$ *to* $m-1$

27. $xtag\leftarrow y\oplus xtoken_i$

28. $r[j]\leftarrow XMap[xtag]$

29. END FOR

30. $res[i]\leftarrow(e,r)$

31.END FOR

定理2 设泄露函数 $\mathcal{L}_{cEMM}(MM,Q)$ 以多映射 MM 与查询序列 Q 为输入,输出 (N, SP, RP),则加密多映射方案 RH-cEMM 在泄露函数 $\mathcal{L}_{cEMM}(MM,Q)$ 下满足定义4所给出的结果隐藏安全性。

证明: 敌手 $\mathcal{A}$ 任意选取两个多映射 MM_1、MM_2,满足以下两个条件:①标签集 $\mathbb{L}$ 大小相等,同为 n;②对任意标签 $l\in\mathbb{L}$,$|MM_0[l]|=|MM_1[l]|$。对于任意查询 φ,RH-cEMM 方案 *XMap* 中的每个值都由伪随机函数的输出加密,满足"一次一密"原则,因此敌手无法通过观察查询结果 $r[j]$ 推断其明文的比特值。换而言之,敌手看到的查询结果分布与 $\{0,1\}$ 均匀分布不可区分,因此敌手 $\mathcal{A}$ 无法通过查询结果区分两个多映射。

定理3 设 Σ 为 $\mathcal{L}$-自适应安全的联合查询加密多映射方案,如果 Σ 满足定义4所给出的结果模式隐藏安全性,那么该方案满足 KPRP 隐藏。

证明: 根据文献[13,16,17]所示,KPRP 指的是一次查询中任意两个标签匹配值集合之间的交集。设查询 $\varphi=l_1\wedge\ldots\wedge l_m$,则 KPRP 可表示为 $MM[l_1]\wedge MM[l_j]$,其中 $2\leqslant j\leqslant m$。由于 Σ 满足定义3所给出的结果模式隐藏安全性,则敌手看到的每个查询结果 $r[j]$ 的分布与 $\{0,1\}$ 均匀分布不可区分,因此敌手无法正确计算 $MM[l_1]\wedge MM[l_j]$,$2\leqslant j\leqslant m$,于是 Σ 满足 KPRP 隐藏。

4 性能分析

4.1 理论性能分析

本文从通信带宽、搜索计算复杂度两个方面对 STE-BQ、vRHS、Doris 与 RH-cEMM 方案进行比较[17-19]。其中,通信带宽开销由客户端发送的搜索令牌、服务器返回的结果集合两部分组成。在搜索令牌方面,RH-cEMM 方案的搜索令牌开销仅与查询标签数 m 线性相关,优于与 $|MM[l_1]|\cdot m$ 相关的 RHS、Doris,而 STE-BQ 的令牌开销为固定常数。在结果返回上,vRHS 与 RH-cEMM 需返回整个 $MM[l_1]$,而 STE-BQ、Doris 仅返回最终结果 $MM[\varphi]$。

在计算复杂度方面,对于客户端而言,vRHS、Doris 方案均需计算 $\mathcal{O}(|MM[l_1]|\cdot m)$ 次生成搜索令牌,STE-BQ、RH-cEMM 仅需要计算 $\mathcal{O}(m)$ 次生成搜索令牌。在服务器方面,vRHS、Doris 与 RH-cEMM 的计算复杂度一致,均为 $\mathcal{O}(|MM[l_1]|\cdot m)$,而 STE-BQ 的计算复杂度为 $\mathcal{O}(|MM[l_1]|)$

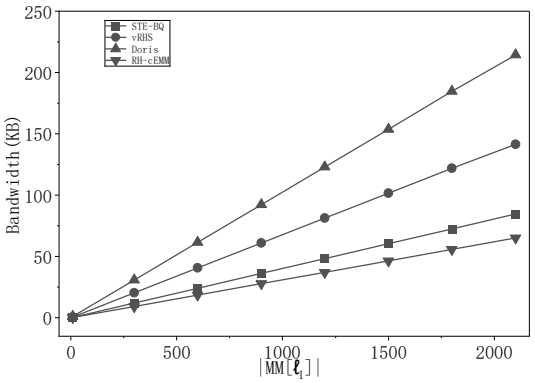
4.2 实验结果

本文测试设备的操作系统为 Ubuntu 22.04,配备了 AMD Ryzen 5 7500F 3.7 GHz 的处理器与 32 GB 内存,实验方案与 TSet 实例均采用 Python 进行实现。同时,方案中的密码学操作采用 Pycrypto 库的方法实现(包括 AES-256、SHA-256、HMAC-SHA256)。具体而言,方案采用 CBC 模式的 AES-256 实现对称加密方案,使用 SHA-256 实现哈希函数,并通过 HMAC-SHA256 实现伪随机函数。

图1为在两标签查询场景下($l_1\wedge l_2$),STE-BQ、RHS、

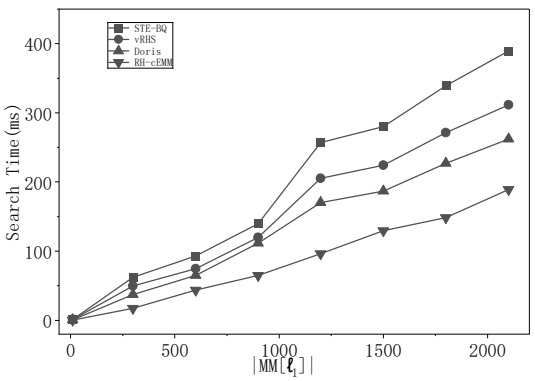
Doris 与本文方案在不同 $|MM[l_1]|$ 时的通信带宽和搜索时间。由此可见，随着 $|MM[l_1]|$ 的逐渐增加，4 种方案的通信开销、搜索时间均呈线性增长。在通信带宽方面，RH-cEMM 通信开销为 $|MM[l_1]|+2$，而 Doris 的通信开销为 $|MM[l_1]|+MM[\varphi]$，显然前者远小于后者（RH-cEMM 相较于 Doris 提升 69.75%），与理论分析一致。在搜索时间方面，RH-cEMM 相较于 Doris 提升 52.7%，原因为 RH-cEMM 仅需计算更少的搜索令牌。

图 2 为 4 种方案在 $|MM[l_1]|=1\,225$ 时，不同标签数量下的通信带宽、搜索时间。由此可见，随着查询标签数量增长，STE-BQ、RHS、Doris 与 RH-cEMM 的通信开销、搜索时间均呈线性增长，RH-cEMM 相较于 Doris 在通信开销方面提升 83.6%，搜索时间提升 23.7%，与理论分析一致，即 Doris 与查询标签数呈“乘性相关”，而 RH-cEMM 与查询标签数呈“加性相关”。



（a） Bandwith

（a） 通信带宽



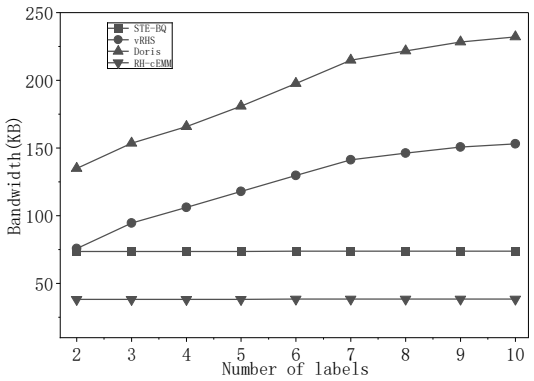
（b） Search time

（b） 搜索时间

Fig. 1 Performance of four schemes under different MM[l_1] lengths

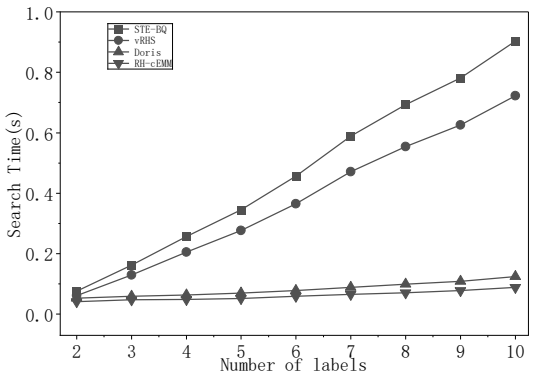
图 1 4 种方案在不同 MM[l_1] 长度下的性能

表 2 为 4 种方案在不同数据集下的服务器存储开销，MM_1、MM_2、MM_3、MM_4 分别包含 10^3、10^4、10^5、10^6 个标签值对。由此可知，所有方案的存储开销均随数据集规模扩大显著增长，但各方案在增长速度上存在明显差异。具体地，RH-cEMM 与 Doris 在小规模数据集下的存储开销较接



（a） Bandwith

（a） 通信带宽



（b） Search time

（b） 搜索时间

Fig. 2 Performance of 4 schemes under different query tag quantities

图 2 4 种方案在不同查询标签数量下的性能

近，但随着数据规模扩大到 MM_4，存储增长明显快于 Doris，原因为 RH-cEMM 为隐藏最终结果模式，编码了任意两个标签与相应值的包含关系而引入了额外的空间成本，但这些额外成本在实际应用场景中是可接受的。

Table 2 Server storage overhead of 4 different solutions

表 2 4 种方案的服务器存储开销 （MB）

方案	MM_1	MM_2	MM_3	MM_4
STE-BQ[18]	3.56	36.82	378.40	3 899.78
vRHS[19]	0.68	6.17	66.21	682.46
Doris[17]	0.72	6.69	86.35	898.33
RH-cEMM	0.78	10.42	101.03	2466.16

5 结语

本文针对联合查询可搜索加密中的结果模式泄露问题，提出具有结果模式隐藏安全性的 RH-cEMM 方案。通过高效令牌派生与基于随机掩码的布尔关系编码，在单轮通信中抑制了关键字对结果模式、跨查询交集结果模式及更强形式的查询结果模式的泄露，在给定泄露函数下提供了更强的隐私保护。

实验表明，RH-cEMM 有效提升了通信效率，相较于

Doris 降低了 69.75%~83.6%,搜索时间也得到了改善。在保持次线性搜索复杂度的同时,实现了对结果模式的隐藏,平衡了安全性与效率。

当前方案主要针对静态数据集,未来可进一步拓展至支持动态更新的可搜索加密系统,并结合前向与后向安全需求,构建更适应实际场景的联合查询方案。

参考文献:

[1] SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]// Proceeding of IEEE Symposium on Security and Privacy, 2000: 44-54.

[2] CHASE M, KAMARA S. Structured encryption and controlled disclosure [C]// Advances in Cryptology- ASIACRYPT, 2010: 577-594.

[3] BOST R. ∑οφος: forward secure searchable encryption[C]// Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016: 1143-1154.

[4] BOST R, MINAUD B, OHRIMENKO O. Forward and backward private searchable encryption from constrained cryptographic primitives[C]// Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017: 1465-1482.

[5] GHAREH C J, PAPADOPOULOS D, PAPAMANTHOU C, et al. New constructions for forward and backward private symmetric searchable encryption[C]// Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018: 1038-1055.

[6] ZUO C, SUN S F, LIU J K, et al. Dynamic searchable symmetric encryption with forward and stronger backward privacy[C]// Computer Security - ESORICS, 2019: 283-303.

[7] XU P, SUSILO W, WANG W, et al. ROSE: robust searchable encryption with forward and backward security[J]. IEEE Transactions on Information Forensics and Security, 2022,17(1): 1115-1130.

[8] LI J, HUANG Y, WEI Y, et al. Searchable symmetric encryption with forward search privacy[J]. IEEE Transactions on Dependable and Secure Computing, 2021, 18(1): 460-474.

[9] SUN S F, STEINFELD R, LAI S Q, et al. Practical non-interactive searchable encryption with forward and backward privacy[C]// Proceedings of the Network and Distributed System Security Symposium, 2021: 2825-2839.

[10] CAI Y R, MA C S, ZHU Y S. Volume-hiding range encrypted multi-map scheme [J]. Journal of Cryptologic Research, 2025, 12 (2): 353-369.
蔡育锐,马昌社,朱泳诗. 规模隐藏的范围加密多映射方案[J]. 密码学报, 2025, 12(2): 353-369.

[11] SONG X F, WANG H Q. A verifiable multi-keyword searchable encryption scheme with forward and backward security[J]. Chinese Journal of Computers, 2023, 46(4): 727-742.
宋翔飞, 王化群. 一个具有前向安全和后向安全的可验证多关键字可搜索加密方案[J]. 计算机学报, 2023, 46(4): 727-742.

[12] ZHANG W Q, LI X, YIN Z, et al. A robust forward and backward private conjunctive symmetric searchable encryption scheme[J]. Journal of Software, 2025, 36(8): 3858-3882.
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.

[13] CASH D, JARECKI S, JUTLA C, et al. Highly-scalable searchable symmetric encryption with support for Boolean queries[C]// Advances in Cryptology-CRYPTO, 2013:353-373.

[14] ISLAM M S, KUZU M, KANTARICIOGLU M. Access pattern disclosure on searchable encryption: Ramification, attack and mitigation[C]// San Diego: 19th Annual Network and Distributed System Security Symposium, 2012.

[15] ZHANG Y P, KATZ J, PAPAMANTHOU C. All your queries are belong to us: the power of file-injection attacks on searchable encryption[C]// Proceedings of the 25th USENIX Conference on Security Symposium, 2016: 707-720.

[16] LAI S Q, PATRANABIS S, SAKZAD A, et al. Result pattern hiding searchable encryption for conjunctive queries[C]// Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018: 745-762.

[17] WANG Y, SUN S F, WANG J F, et al. Practical non-interactive encrypted conjunctive search with leakage suppression[C]// Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security, 2024: 4658-4672.

[18] CHEN L, MU Y, YANG J, et al. Keyword-pair result pattern hiding structured encryption for Boolean queries[J]. IEEE Transactions on Information Forensics and Security, 2025, 20(1): 8449-8461.

[19] DENG Q, CHEN L, ZHU Y, et al. Leakage reduced searchable symmetric encryption for multi-keyword queries[J]. IEEE Transactions on Cloud Computing, 2025, 13(3): 882-894.

[20] 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.

[21] GOH E. Secure indexes[EB/OL]. https://eprint. iacr. org/2003/216. pdf.

[22] KAMARA S, MOATAZ T. Boolean searchable symmetric encryption with worst-case sub-linear complexity[C]// Advances in Cryptology-EUROCRYPT, 2017: 94-124.

[23] PATEL S, PERSIANO G, SEO J Y, et al. Efficient Boolean search over encrypted data with reduced leakage[C]// Advances in Cryptology-ASIACRYPT, 2021: 577-607.

[24] BELLARE M, DESAI A, JOKIPII E, et al. A concrete security treatment of symmetric encryption[C]// Proceedings of the 38th Annual Symposium on Foundations of Computer Science, 1997: 394-403.

[25] BELLARE M, CANETTI R, KRAWCZYK H. Pseudorandom functions revisited: The cascade construction and its concrete security[C]// Proceedings of the 37th Conference on Foundations of Computer Science, 1996: 514-523.

(责任编辑:刘嘉文)