抗诬陷码及相关组合结构的界与算法

Marco Dalai*、Stefano Della Fiore*†、Adele A. Rescigno† 和 Ugo Vaccaro†
*布雷西亚大学信息工程系，布雷西亚（BS），意大利 电子邮箱：{marco.dalai、s.dellafiore001}@unibs.it
†萨莱诺大学计算机科学系，菲夏诺（SA），意大利 电子邮箱：{sdellafiore、arescigno、uvaccaro}@unisa.it

arXiv:2303.07211v1 [cs.IT] 13 Mar 2023

摘要——本文研究了 Boneh 和 Shaw[3] 提出的、用于保护受版权保护的材料的抗诬陷码的最小长度上界。长度为 t 的 q 元 (k,n)-抗诬陷码是一个 $t\times n$ 矩阵，其元素属于 $\{0,1,\ldots,q-1\}$，并满足以下性质：对于任意一列 c 以及另外的 k 列，存在某一行，使得 k 列中的符号均不同于该行中列 c 的对应符号。本文利用 Lovász 局部引理，证明了存在长度为 $t=O(\frac{k^2}{q}\log n)$ 的 q 元 (k,n)-抗诬陷码（当 $q\le k$ 时）；并利用删余法，证明了存在长度为 $t=O(^{k}{}_{\log(q/k)}\log(n/k))$ 的此类码（当 $q>k$ 时）。值得注意的是，在 $q\le k$ 的实际情形下，我们的结果给出了长度几乎达到任意 q 元 (k,n)-抗诬陷码长度下界 $\Omega(_{q}{}^{k^2}{}_{\log k}\log n)$ 的码；更重要的是，我们还得到了一个复杂度为 $O(tn^2)$ 的算法，用于构造此类码。

索引词——抗诬陷码、强选择码、叠加码、Lov´asz 局部引理。

I. 引言

抗诬陷码由 Boneh 和 Shaw [3]提出，旨在保护受版权保护的材料免遭未经授权的使用。其原理如下。当分销商想要销售数字产品的 n 份副本时，她首先在数字产品中选定 t 个固定位置。随后，她为每份产品副本中的每个选定位置关联一个 q元符号。每份产品副本中这些选定符号及其位置的集合称为指纹，可视为一个长度为 t、字母表大小为 q的码字。客户不知道数据中存储的符号及其位置，因此无法将其删除或修改。然而，一群恶意客户可能会串通起来，分享并比对各自的副本。在这种情况下，恶意客户原则上可以推断出指纹的符号及其位置，从而制作非法副本，也就是指纹与恶意合谋联盟之外某个无辜客户的指纹相同的数字产品副本。为防止这种情况，可以使用抗诬陷码。非正式地说，如果一组指纹（码字）构成一个 (k,n)-抗诬陷码，那么任何由至多k 名客户组成的合谋联盟都无法生成联盟之外的其他码字。从本质上说，这意味着每个码字都包含一条信息（位于某个坐标上），而任何至多由 k 个其他码字组成的集合，在相同坐标上都不包含这条信息（更正式的定义将在后文给出）。

该领域的主要问题包括：给出抗诬陷码长度的界（其动机是实际应用中的考量：嵌入数字数据中的指纹长度会带来显而易见的额外开销，因此需要尽可能缩短），以及设计计算效率高的方法，构造长度接近理论最优值的抗诬陷码。开创性论文 [3] 推动了对上述两个问题的深入研究。由于篇幅所限，此处无法概述该领域的诸多研究成果，相关文献可参阅论文 [1], [2], [4], [5], [10], [15]–[20] 及其中引用的参考文献。

A. 我们的结果

本文利用 Lovasz 局部引理证明，存在长度为 $t=O(n)$ 的 q 元 (k,n)-抗诬陷码（适用于 $q\le k$），也存在长度为 $t=(\frac{k^2}{q}\overline{\log})$ 的 q 元 (k,n)-抗诬陷码（适用于 $q>k$），后者通过删余法构造。类似地，我们请读者参阅 [8] 和 [11]，其中运用这些方法对分离码和可识别父代性质（IPP）码给出了概率构造。证明结果之后，我们将讨论这些结果如何改进文献中的已知结果。此外，对于实际中重要的情形 $q\le k$（其动机见 [15]），我们给出一个 $O(tn^2)$ 随机化算法，用于构造长度为 $t=O(\frac{k^2}{q}\log n)$ 的码；该长度几乎达到任意 q 元 (k,n)-抗诬陷码长度下界 $\Omega(_{q}{}^{k^2}{}_{\log k}\log n)$。据我们所知，这是首个关于码大小 n、具有此类性能的多项式时间算法。注意，显式构造长度为 t 的 (k,n)-抗诬陷码，任何算法都需要 $\Omega(tn)$ 时间。

II. 预备知识

本文中，不带下标的对数均以 2 为底，我们用 $\ln(\cdot)$ 表示自然对数。给定整数 $a<b$，我们用 $[a,b]$ 表示集合 $\{a,a+1,\ldots,b\}$。首先介绍本文研究的组合对象。

定义 II.1. 设 k、n、$q\ge 2$ 为正整数，$n>k$。一个 q 元 (k,n)-防伪编码是一个 $t\times n$ 矩阵 M，其元素属于 $[0,q-1]$，并且对于任意一列 c 及任意其他

k 列 M，都有一行 $i \in [1,t]$，使得这 k 列中的符号均不同于列 c 的第 i 行中的符号。M 的行数 t 称为 q 元 (k,n)-防伪编码的长度。

为了给出 q 元 (k,n) 抗诬陷码的最小长度上界，我们需要回顾一类与之密切相关的码，即 q 元 (k,n) 强选择码，该类码由 De Bonis 和 Vaccaro 在 [7] 中研究。

定义 II.2. 设 k、n 和 $q \geq 2$ 为正整数，$n \geq k$。q 元 (k,n) 强选择码是一个 $t \times n$ 矩阵 M，其元素取自 $[0, q-1]$，并满足：对于 M 的任意 k 元列组，以及给定 k 元列组中的任意一列 c，都存在一行 $i \in [1,t]$，使得 c 在第 i 行的元素为 $s \in [1, q-1]$，而该 k 元列组中其余所有列在第 i 行的元素均属于 $[0, q-1] \setminus \{s\}$。M 的行数 t 称为 q 元 (k,n) 强选择码的长度。

给定整数 $q \geq 2$ 和一个 q 元向量 $\mathbf{c} \in [0, q-1]^t$，我们用 $w(\mathbf{c})$ 表示 c 的非零分量数。

定义 II.3. q 元 (k,w,n)-强选择码是满足附加条件的 q 元 (k,n)-强选择码：每一列 c 都有 $w(\mathbf{c})$ 个 $= w$。

q 元 (k,n)-抗诬陷码与 q 元 $(k+1,n)$-强选择码之间存在紧密联系。事实上，若以 $t_{\mathrm{FP}}(q,k,n)$ 表示 q 元 (k,n)-抗诬陷码的最小长度，并以 $t_{\mathrm{SS}}(q,k,n)$ 表示 q 元 (k,n)-强选择码的最小长度，则有以下引理。

引理 II.4。

$$t_{SS}(q,k+1,n) \geq t_{FP}(q,k,n) \geq \frac{1}{2} t_{SS}(q,k+1,n)\,.$$

证明。注意到，根据这类码的定义，长度为t 的 q 元 $(k+1, n)$-强选择码也是长度为 t 的 q 元 (k,n)-抗诬陷码，因此 $t_{\mathrm{FP}}(q,k,n)$ 的上界很容易得到。反过来，q 元 (k,n)-抗诬陷码 $\mathcal{C}$ 不一定是 q 元 $(k+1,n)$-强选择码，因为可能存在一列 c 以及 $\mathcal{C}$ 中其他 k 列，使得其中仅有的那些行，其k 列的符号不包含列 c 的符号，并且该符号等于 0。因此，这满足抗诬陷码的性质，却不满足强选择码的性质。然而，q 元 $(k+1,n)$-强选择码的最小长度 $t_{\mathrm{SS}}(q,k+1,n)$ 至多是 q 元 (k,n)-抗诬陷码的最小长度 $t_{\mathrm{FP}}(q,k,n)$ 的两倍。事实上，给定一个 q 元 (k,n)-抗诬陷码 $\mathcal{C}$，可以通过取 $\mathcal{C}$ 的行与互补码 $\overline{\mathcal{C}}$ 的行的并集，构造一个长度为 $\mathcal{C}$ 两倍的 $(k+1,n)$-强选择码；互补码 $\overline{\mathcal{C}}$ 是将 $\mathcal{C}$ 中每个元素 s 替换为 $q-1-s$ 得到的。因此，我们得到引理陈述中所示的$t_{\mathrm{FP}}(q,k,n)$下界。

□

根据引理 II.4，我们可以集中研究如何求得最小长度 $t_{\mathrm{SS}}(q,k,n)$ 的界，进而间接得到 $t_{\mathrm{FP}}(q,k,n)$ 的界。

III. 利用洛夫什兹局部引理构造抗诬陷码的新随机化算法

本节给出 q 元（k,n）强选择码的最小长度 $t_{\mathrm{SS}}(q,k,n)$ 的新上界。我们将给出两种不同的界：一种利用 Lov´asz 局部引理推导，另一种则利用删余法。

文献 [7]，已证明存在长度为

$$t = O\left(\frac{k^2}{v}\log(n/k)\right), \tag{1}$$

其中 $v = q-1$（当 $q \leq k$ 时），且 $v = k$（当 $q > k$ 时）。他们的概率构造给出了一种复杂度为 $\Theta(n^k)$ 的随机化算法，用于生成 q 元的（k,n）强选择性码，其长度与公式 (1) 中给出的长度同阶。对于较小的 k 值，该算法显然已不切实际。本文给出了关于 $t_{\mathrm{SS}}(q,k,n)$ 的新界，其阶与 [7] 给出的界相同，但关键在于，同时还给出了一种复杂度为 $O(tn^2)$ 的随机化构造算法。这意味着，我们可以在关于 n 和 k 的多项式时间内，构造长度如 (1) 所示的 q 元（k,n）强选择性码。

关键思路是使用以下矩阵（在 [12] 中引入），其中约束仅涉及成对的列。

定义 III.1. 设 n、w、λ 为正整数。若一个 q 元 $t \times n$ 矩阵 M 的元素取自 $[0, q-1]$，且满足以下性质，则称其为 (λ, w, n)-矩阵：

1) M 的每一列 c 都有 w 个 $= w$，
2) M 的任意两列 c、d 至多有 λ 个相同的非零符号；也就是说，至多存在 λ 行，使得列 c 和 d 的对应元素均为 $s \in [1, q-1]$。

这些矩阵与 q 元 (k,w,n)-强选择码之间的关系由以下结果给出。

引理 III.2. 若 M 是一个 q 元 $t \times n$ (λ, w, n)-矩阵，且满足 $\lambda = \lfloor (w-1)/(k-1) \rfloor$，则 M 是一个长度为 t 的 q 元 (k,w,n)-强选择码。

证明。设 c 是 M 的任意一列，并令 A 为列 c 中非零元素所在的行索引集合。因此 $|A| = w$。令 K 为 M 中任意一组 $k-1$ 列，且 c $\notin K$。将 $M(A,K)$ 记为 M 的 $w \times (k-1)$ 子矩阵：先选取 K 中的 $k-1$ 列，再选取行索引属于 A 的 w 行。由于 M 是一个 (λ, w, n)-矩阵，因此列 c 与 K 中任意一列在相同行中共有的非零元素至多为 $\lambda = \left\lfloor \frac{w-1}{k-1} \right\rfloor$。因此，$M(A,K)$ 中非零元素的总数至多为

$$(k-1)\lambda \leq (k-1)\left\lfloor \frac{w-1}{k-1} \right\rfloor \leq w-1\,.$$

鉴于 $M(A,K)$ 具有 $|A| = w$，我们可知 $M(A,K)$ 中至少有一行仅包含零元素。因此引理得证。

□

现在，我们想为 q 元 (λ, w, n) 矩阵的最小长度给出一个较好的上界，从而为 q 元 (k, w, n) 强选择码的最小长度提供一个上界。首先，我们需要回顾以下关于二项式系数的众所周知的事实：对于正整数 $c \le b \le a$，有

$$\left(\frac{a}{b}\right)^b \le \binom{a}{b} \le \frac{a^b}{b!} \le \left(\frac{ea}{b}\right)^b, \tag{2}$$

$$\binom{a}{b}\binom{b}{c} = \binom{a}{c}\binom{a-c}{b-c}, \tag{3}$$

以及下面这个有用的不等式，其证明见 [9], , 适用于正整数 $c \le a \le b$

$$\binom{a}{c} \Big/ \binom{b}{c} \le \left(\frac{a - \frac{c-1}{2}}{b - \frac{c-1}{2}}\right)^c. \tag{4}$$

我们的主要工具是著名的 Lov´asz 局部引理的对称情形算法版本，见 [13]。

引理 III.3. *[13]* 设 $E_1, E_2, \ldots, E_m$ 是概率空间中的事件，其中每个事件 E_i 都与其余所有事件 E_j 组成的集合相互独立，但至多有 D 个例外，并且对所有 $1 \le i \le m$，$\Pr(E_i) \le P$。若 $ePD \le 1$，则 $\Pr(\cap_{i=1}^m \overline{E_i}) > 0$。此外，可以通过使用平均不超过 m/D 次重新抽样，有效地找到一个避开所有事件 E_i 的配置。

现在我们可以陈述主要引理了。

引理 III.4. 存在一个 q 元 $t \times n$ (λ, w, n) 矩阵，其

$$t = \max\left\{ \lceil 2w - (\lambda+1) \rceil, \left\lceil \frac{\lambda}{2} + \frac{1}{q-1}\left(\frac{ew}{\lambda+1}\left(w - \frac{\lambda}{2}\right)(e(2n-4))^{\frac{1}{\lambda+1}}\right)\right\rceil \right\}. \tag{5}$$

证明。令 M 为一个随机 $t \times n$ q 元矩阵，$t \ge 2w - (\lambda+1)$，其中每一列 $\mathbf{c}$ 都从所有满足 $w(\mathbf{c}) = w$ 的长度为 t 的互异 q 元向量 $\mathbf{c}$ 构成的集合中均匀随机选取。容易看出，此类向量的数量等于 $\binom{t}{w}(q-1)^w$，因此

$$\Pr(\mathbf{c}) = \left(\binom{t}{w}(q-1)^w\right)^{-1}.$$

设 $i, j \in [1, n], i \ne j$，并考虑事件 $\overline{E}_{i,\,j}$：存在至多 λ 行，使得在这些行中，M 的第 i 列和第 j 列均含有相同的非零符号。我们评估其补“坏”事件 $E_{i,\,j}$ 的概率。因此，$E_{i,\,j}$ 表示第 i 列和第 j 列 $\mathbf{c}_i$ 与 $\mathbf{c}j$ 在至少 $\lambda+1$ 个坐标上具有相同的非零元素这一事件。我们通过对事件 $\mathbf{c}_i$ 等于 c 进行条件化，来界定 $\Pr(E_{i,\,j})$。

对于坐标子集 $S \subset [1, t]$，令 $E^S_{i,\,j}$ 表示这样一个事件：在 S 的每个坐标中，第 i列和第 j列具有相同的非零元素。最后，令 A 为.....的集合

其中 $\mathbf{c}_i$ 非零的坐标。注意，对于 $S \in \binom{A}{\lambda+1}$，即 S 是 A 的一个大小为 $\lambda+1$的子集，我们有

$$\Pr(E^S_{i,j} | \mathbf{c}_i = c) = \frac{\binom{t-(\lambda+1)}{w-(\lambda+1)}(q-1)^{w-(\lambda+1)}}{\binom{t}{w}(q-1)^w}.$$

Then

$$\begin{aligned}\Pr(E_{i,j} | \mathbf{c}_i = c) &\le \sum_{S \in \binom{A}{\lambda+1}} \Pr(E^S_{i,j} | \mathbf{c}_i = c) \\ &= \binom{w}{\lambda+1} \frac{\binom{t-(\lambda+1)}{w-(\lambda+1)}(q-1)^{w-(\lambda+1)}}{\binom{t}{w}(q-1)^w}. \end{aligned} \tag{6}$$

由于 (6) 的右端不依赖于固定列 c，因此该结论无条件时也成立。因此

$$\Pr(E_{i,j}) \le \binom{w}{\lambda+1} \frac{\binom{t-(\lambda+1)}{w-(\lambda+1)}(q-1)^{w-(\lambda+1)}}{\binom{t}{w}(q-1)^w}. \tag{7}$$

因此，由 (7) 可得

$$\begin{aligned}\Pr(E_{i,j}) &\le \binom{w}{\lambda+1}\binom{t-(\lambda+1)}{w-(\lambda+1)} \Big/ \binom{t}{w}(q-1)^{\lambda+1} \\ &\overset{(i)}{=} \binom{w}{\lambda+1}\binom{w}{\lambda+1} \Big/ \binom{t}{\lambda+1}(q-1)^{\lambda+1} \\ &\overset{(ii)}{\le} \frac{1}{(q-1)^{\lambda+1}}\binom{w}{\lambda+1}\left(\frac{w - \frac{\lambda}{2}}{t - \frac{\lambda}{2}}\right)^{\lambda+1} \\ &\overset{(iii)}{\le} \frac{1}{(q-1)^{\lambda+1}}\left(\frac{ew}{\lambda+1}\right)^{\lambda+1}\left(\frac{w - \frac{\lambda}{2}}{t - \frac{\lambda}{2}}\right)^{\lambda+1} = P, \end{aligned} \tag{8}$$

其中，(i) 根据等式 (3) 成立，(ii) 根据不等式 (4) 成立，最后，(iii) 根据不等式 (2) 成立。

事件 $E_{i,\,j}$ 的数量等于 $n(n-1)/2$。固定一个事件 $E_{i,\,j}$，依赖于该事件的事件 $E_{i,\,j}$ 的数量等于 $D = 2n-4$。因此，根据引理 III.3，若 P（定义见 (8)）和 $D = 2n-4$ 满足 $ePD \le 1$，则所有“坏”事件 $E_{i,\,j}$ 均不发生的概率严格大于零。可以看出，将 t 设为 (5) 中最大值的第二项，确实可得

$$ePD = \frac{e(2n-4)}{(q-1)^{\lambda+1}}\left(\frac{ew}{\lambda+1}\right)^{\lambda+1}\left(\frac{w - \frac{\lambda}{2}}{t - \frac{\lambda}{2}}\right)^{\lambda+1} \le 1.$$

因此，由引理 III.3 可构造一个 q 元的（λ, w, n）矩阵 M，其行数 t 满足等式 (5)，因为我们还需要考虑初始假设 $t \ge 2w - (\lambda+1)$，以确保本引理中的所有计算都有意义。

□

现在，借助引理 III.2 和引理 III.4，我们可以证明以下结果。

定理 III.5. 存在一种随机化算法，可构造长度为

$$t \le 1 + \max\left\{2w - \frac{w-1}{k-1},\ \frac{w-1}{2(k-1)} +\right.$$

$$\frac{ew(k-1)}{(q-1)(w-1)}\left(w-\frac{w-1}{2(k-1)}+\frac{1}{2}\right)(e(2n-4))^{\frac{k-1}{w-1}}\Bigg\}. \tag{9}$$

该算法构造此编码平均需要 $O(tn^2)$ 时间。

证明。将引理 III.2 中 $\lambda = \lfloor (w-1)/(k-1) \rfloor$ 的值代入引理 III.4 的方程（5），并利用不等式 $\frac{w-1}{k-1} - 1 \leq \left\lfloor \frac{w-1}{k-1} \right\rfloor \leq \frac{w-1}{k-1}$，即可得出引理陈述中关于 t 的上界。时间复杂度 $O(tn^2)$ 由引理 III.3 得出，注意到 $m/D = n(n-1)/(4n-8) \leq n/3$。此外，根据 [13] 得到的算法需要随机生成一个矩阵，检查是否满足 $\Theta(n^2)$ 个事件 $\overline{E}_{i,\,j}$，并且仅对未满足的事件进行重新抽样。这意味着我们需要检查矩阵的第 i 列和第 j 列是否至多有 $\left\lfloor \frac{w-1}{k-1} \right\rfloor$ 个非零元素重合；这至多需要 $O(t)$ 次运算，然后仅对未满足的事件重新抽样。接着，我们只需检查涉及已重新抽样列的事件。总之，根据引理 III.3，该过程需要 $O(tn^2 + n \cdot m/D \cdot t) = O(tn^2)$ 次基本运算。

□

注 III.6. 在对参数 w进行优化（9）之前，我们想强调，定理 III.5 并非单纯的技术性中间结果，在若干实际场景中也可能具有重要意义。例如，已有研究表明，q元（k,n）强选择性码可用于解决多信道无线网络中出现的重要通信问题。在 *[7],* 所考虑的场景中，有一组 n 个非协调站点，尝试通过一组 $q-1$ 个独立信道进行传输。仅当没有两个站点在同一时刻尝试通过同一信道传输时，传输才会成功。*[7]* 中的思路如下：为每个站点分配一个 q元（k,n）强选择性码中的不同码字，每个码字对应于相关站点的传输时序表。给定码字 c 的第 i个坐标中出现符号 $s \in [0, q-1]$，自然可理解为向持有码字 c 的站点发出“指令”：若$s = 0$，则在通信协议的第 i步保持“静默”；若 $s = j \neq 0$，则在第 i步通过第 j个信道传输。在任意给定时刻至多有k 个站点处于“活动”状态这一假设下，*[7]* 的作者证明了，q元（k,n）强选择性码自然对应于多信道无线网络中的冲突解决协议。如今，在许多情况下，不仅要尽可能缩短协议长度（即所有站点成功传输之前的时刻数），还要限制每个站点尝试传输的次数（例如为了节省能量；更多内容见 *[9]* ）。因此，在这些情况下，q元（k,w,n）强选择性码可能会派上用场，其中每个码字都有 w 个非零分量。我们指出，*[7]* 中的技术既无法处理对每个站点尝试传输次数的这一严格限制，也无法像我们的方法那样，为构造 w元（ *[7]* , q）强选择性码提出高效算法

每个站点的传输次数，也无法像我们的方法那样，为构造 q元（k,n）强选择性码提出高效算法。

我们优化方程 (9) 中的 w，以获得一种用于（无约束）q 元（k,n）强选择码的随机化算法。

推论 III.7. 设 $w = \lceil 1 + (k-1)\ln(2en) \rceil$，则定理 III.5 中描述的算法构造出一个 q 元（k,w,n）强选择码；显然，它也是一个长度 t 上界如下的 q 元（k,n）强选择码

$$t \leq \max\left\{2(k-1)\ln(2en) - \ln(n),\ \frac{\ln(n)}{2} + \frac{e^2(k-1)^2}{q-1}\ln(2en) + \frac{7e^2(k-1)}{2(q-1)}\right\} + O(1).$$

证明. 固定推论陈述中的 w，并利用以下不等式

$$1+(k-1)\ln(2en) \leq \lceil 1+(k-1)\ln(2en) \rceil \leq 2+(k-1)\ln(2en)$$

根据定理 III.5，只考虑最大值中的第二项，可得

$$\begin{aligned} t &\leq 1 + \frac{1+(k-1)\ln(2en)}{2(k-1)} + \frac{e}{(q-1)\ln(2en)} \\ &\quad (2+(k-1)\ln(2en))\left(\frac{3}{2}+(k-1)\ln(2en)\right)(2en)^{\frac{1}{\ln(2en)}} \\ &\leq \frac{\ln(n)}{2} + \frac{e^2(k-1)^2}{q-1}\ln(2en) + \frac{7e^2(k-1)}{2(q-1)} + O(1), \end{aligned}$$

由于 $k \geq 2$，$\ln(2en) \geq 2$ 且 $(2en)^{\frac{1}{\ln(2en)}} = e$。接着，考虑定理 III.5 中最大值的第一项后，即可得出推论。

□

因此，根据引理 II.4 和推论 III.7，我们得到了本文的主要结果。

定理 III.8. 存在一种随机化算法，可构造长度为

$$t \leq \max\left\{2k\ln(2en) - \ln(n),\ \frac{\ln(n)}{2} + \frac{e^2k^2}{q-1}\ln(2en) + \frac{7e^2k}{2(q-1)}\right\} + O(1).$$

该算法平均需要 $O(tn^2)$ 的时间来构造该编码。

为了恰当地评估定理 III.8 的价值，我们回顾以下结果，该结果给出了任意 q-元 (k,n)-防伪编码长度的下界。

定理 III.9. *[6], [7], [15]* 给定正整数 q、k 和 n，且满足 $q \geq 2$ 和 $2 \leq k \leq \sqrt{n}$，任何 q 元 (k,n)-防伪码的最小长度满足

$$t_{FP}(q,k,n) = \Omega\left(\frac{k^2}{q\log k}\log\frac{n}{k}\right). \tag{10}$$

因此可以看出，定理 III.8 提供的构造方法不仅相当高效，而且能生成长度几乎最优的码。

我们还可以证明以下结果。

定理 III.10. 设 M 是一个最小长度为 t 的 q 元 (k,n)-防伪编码。则

$$\left\lceil \frac{n}{q-1} \right\rceil \geq t \geq \left\lceil \frac{1}{q} \min\left\{ n, \frac{15+\sqrt{33}}{24} k^2 \right\} \right\rceil . \qquad (11)$$

证明。我们可以定义一个从 M 到长度为 t' 的二元 (k,n)-强选择编码的映射：将每个符号 $i \in [0, q-1]$ 映射为 e_{i+1}，其中 e_i 是一个二元列向量，其第 i 个分量为 1，其余分量为 0。因此 $t' = qt$。现在，由于 $t' \geq \min\left\{ n, \frac{15+\sqrt{33}}{24} k^2 \right\}$，根据 [14, 定理 2]，即可得到 (11) 的右侧不等式。为证明 (11) 的左侧不等式，只需观察：从并排拼接的 $(q-1)$ 个 $\lceil n/(q-1) \rceil \times \lceil n/(q-1) \rceil$ 对角矩阵中任取 n 列，其中第 i 个矩阵的对角线上为符号 $i \in [1, q-1]$，其余位置为 0，即可得到一个长度为 $\lceil n/(q-1) \rceil$ 的 q 元 (k,n)-防伪编码。

□

IV. 通过删余法改进 $t_{\mathrm{FP}}(q, k, n)$ 的新上界

本节将给出 q 元 (k,n) 抗诬陷码最小长度的一个存在性上界，该上界优于文献中已知的最佳结果。我们首先回顾这些已知结果。

定理 IV.1. *[19]* 存在一个长度为以下数值的 q 元 (k,n) 抗诬陷码：

$$t \leq -k \ln\left(n \frac{k!}{k!-1} \right) \Big/ \ln\left(1 - \left(1 - \frac{1}{q} \right)^k \right) .$$

我们注意到，在原论文 [19]，中，这个上界是以分离哈希族长度的上界形式给出的。然而，众所周知，类型为 $(t, n, q, \{1, k\})$ 的分离哈希族等价于一个长度为 t 的 q 元 (k,n) 抗诬陷码。在 [15]，中，作者给出了以下上界；当 $q \leq k$ 且 k 足够大时，该上界优于定理 IV.1 中的上界。

定理 IV.2. *[15]* 若 $q \leq k$，则存在一个长度为以下数值的 q 元 (k,n) 抗诬陷码

$$t \leq \frac{-k \ln n - (k+1) \ln 2}{\ln\left[1 - \left(1 - \frac{q-1}{k+1} \right) \left(\frac{q-1}{k+1} \right)^k - \frac{q-1}{k+1} \left(1 - \frac{1}{k+1} \right)^k \right]} .$$

我们的结果采用所谓的删余法或删除法，这是 [19]，所用的同一种技术，但对参数进行了更精细的优化。因此，对于参数 k、q 和 n 的所有取值，我们都能改进定理 IV.1 的界。下面陈述以下定理。

定理 IV.3. 存在一个长度为 t 的 (k,n)-无框架编码，其中 t 是满足以下不等式的最小整数

$$(k+1) \binom{n\left(1+\frac{1}{k}\right)}{k} (1 - p_{q,k})^t \leq 1 ,$$

其中

$$p_{q,k} := \begin{cases} \left(1 - \frac{1}{q}\right)^k & \text{for } q > k, \\ \left(1 - \frac{q-1}{k+1}\right) \left(\frac{q-1}{k+1}\right)^k + \frac{q-1}{k+1} \left(1 - \frac{1}{k+1}\right)^k & \text{otherwise.} \end{cases}$$

证明。设 M 是一个 q 元 $t \times (n+\ell)$ 矩阵，其中每个元素都从集合 $[0, q-1]$ 中独立同分布地随机选取，其分布为 $\mu = (\mu_0, \mu_1, \ldots, \mu_{q-1})$，该分布将在后文确定。给定索引 $i \in [1, n]$ 和列索引集合 B, $|B| = k-1$, $i \notin B$，令 $E_{i,B}$ 表示如下事件：对于每一行，只要 c_i（第 i 列）包含符号 s，就存在一个索引 $j \in B$，使得 c_j 在同一行中包含符号 s。因此，每个事件 $E_{i,B}$ 的概率可上界为 $\Pr(E_{i,B}) \leq \left(1 - \sum_{i=0}^{q-1} \mu_i (1-\mu_i)^k\right)^t$。此类事件的数量等于 $(n+\ell)\binom{n+\ell-1}{k}$。现在，令 X 为表示满足的事件 $E_{i,B}$ 数量的随机变量。因此，当 $q > k$ 时，对每个 $i \in [0, q-1]$ 取 $\mu_i = 1/q$（均匀分布）；当 $q \leq k$ 时，对每个 $i \in [1, q-1]$ 取 $\mu_i = 1/(k+1)$，并令 $\mu_0 = 1 - (q-1)/(k+1)$，可得

$$\mathbb{E}[X] \leq (n+\ell) \binom{n+\ell-1}{k} (1 - p_{q,k})^t .$$

我们注意到，如果 $\mathbb{E}[X] < \ell + 1$，则至多有 ℓ 个“坏”事件 $E_{i,\,B}$ 成立。然后，对于这些事件中的每一个 $E_{i,\,B}$ 我们移除一列，其索引属于 $\{i\} \cup B$。因此，剩下一个 q 元矩阵，它有 t 行和至少 n 列，且满足抗诬陷性质。因此，我们得到一个长度为 t 的 (k,n)-抗诬陷码。取 $\ell = \lfloor n/k \rfloor$，定理即得证。 □

推论 IV.4. 利用公式 (2) 中的不等式，由定理 IV.3 可得，(k,n)-抗诬陷码的长度上界如下：

$$t \leq \frac{-k \ln\left(n \frac{k+1}{k}\right) - \ln\left(\frac{k+1}{k!}\right)}{\ln(1 - p_{q,k})} ,$$

其中 $p_{q,\,k}$ 是定理 IV.3 中定义的同一量。

定理 IV.5. 对于任意 $n \geq k \geq 2$ 和 $q > k$，推论 IV.4 的界都优于定理 IV.1 中的界。

证明。我们需要证明以下不等式

$$\frac{-k \ln\left(n \frac{k+1}{k}\right) - \ln\left(\frac{k+1}{k!}\right)}{\ln(1 - p_{q,k})} < \frac{-k \ln\left(n \frac{k!}{k!-1}\right)}{\ln\left(1 - \left(1 - \frac{1}{q}\right)^k\right)} . \qquad (12)$$

由于对于 $q > k$, $p_{q,k} = (1 - 1/q)^k$，我们可以重新排列并简化（12）中的各项，从而得到以下不等式。

$$\left(\frac{k+1}{k}\right)^k \frac{k+1}{k!} < \left(\frac{k!}{k!-1}\right)^k . \qquad (13)$$

现在，由于 $\left(1+\frac{1}{k}\right)^k \le e$ 和 $k! \ge e\left(\frac{k}{e}\right)^k$，（13）的左侧小于 $\left(\frac{e}{k}\right)^k (k+1)$。要证明不等式（13），只需证明

$$\left(\frac{e}{k}\right)^k (k+1) < 1\,, \tag{14}$$

由于对于每个 k，(13) 的右端都大于 1。

可以看出，对于 $k \ge 2$，(14) 的左端是关于 k 的递减函数。使不等式 (14) 成立的第一个整数 k 是 $k = 5$。然后，由于通过直接计算可知，不等式 (13) 对 $k = 2$、3、4 也成立，因此定理得证。□

定理 IV.6. 推论 IV.4 的界对于每个 $n \ge k \ge 2$ 和 $q \le k$ 都优于定理 IV.2 的界。

证明. 显然，我们需要证明以下不等式

$$\frac{-k\ln\left(n\frac{k+1}{k}\right) - \ln\left(\frac{k+1}{k!}\right)}{\ln\left(1-p_{q,k}\right)} < \frac{-k\ln n - (k+1)\ln 2}{\ln\left[1-\left(1-\frac{q-1}{k+1}\right)\left(\frac{q-1}{k+1}\right)^k - \frac{q-1}{k+1}\left(1-\frac{1}{k+1}\right)^k\right]}\,. \tag{15}$$

对于 $q \le k$，根据 $p_{q,\,k}$ 的定义，(15) 中的分母相等。因此，我们可以重新排列并简化各项，得到以下不等式。

$$\left(\frac{k+1}{k}\right)^k \frac{k+1}{k!} < 2^{k+1}\,. \tag{16}$$

按照定理 IV.5 的证明中的方法，对于 $k \ge 2$，(16) 的左端是关于 k 的递减函数，而 (16) 的右端则关于 k 递增。通过观察，很容易看出，即使对于 $k = 2$，不等式 (16) 也成立。□

致谢

A. A. Rescigno 和 U. Vaccaro 的工作部分得到了 SERICS 项目（PE00000014）的支持。该项目属于由欧盟—下一代欧盟资助的 NRRP MUR 计划。

参考文献

[1] A. Barg、G. R. Blakley、G. A. Kabatiansky，《数字指纹编码：问题陈述、构造与叛徒识别》，IEEE 信息论汇刊 49(4)，第 852–865 页，2003 年。[2] S.R. Blackburn，《抗诬陷码》，SIAM 离散数学杂志，13(3)，第 499–510 页，2003 年。[3] D. Boneh 和 J. Shaw，《数字数据的抗合谋指纹》，IEEE 信息论汇刊 44(5)，第 1897–1905 页，1998 年。[4] Y. Chee 和 X. Zhang，《抗诬陷码的改进构造》，IEEE 信息论汇刊 58(8)，第 5449–5453 页，2012 年。[5] M. Cheng 和 Y. Miao，《多媒体指纹中的抗合谋码与检测算法》，IEEE 信息论汇刊，57 7 4843 4851, 2011, ( )，第 - A. De Bonis 和 U. Vaccaro，[6]《一种新型选择器及其在无线多信道网络冲突解决中的应用》。载于：Chrobak, M., Fern´andez Anta, A., Gasieniec, L., Klasing, R. (编)《传感器系统算法》。ALGOSENSORS 2016。计算机科学讲义，第 10050 卷。Springer，2016 年。[7] A. De Bonis 和 U. Vaccaro，《一种新型选择器及其在无线多信道网络冲突解决中的应用》，《理论计算机科学》806，第 219–235 页，2020 年。[8] M. Fernandez 和 J. Livieratos，《分离码构造的算法方面》，《实验算法分析：特别事件》，Springer 国际出版，2019 年。

[9] L. Gargano、A. A. Rescigno 和 U. Vaccaro，《低重量叠加码及相关组合结构：界与应用》，《理论计算机科学》806，第 655–672 页，2020 年。[10] C. Guo、D. R. Stinson 和 T. Van Trung，《二元抗诬陷码的紧界》，《设计、编码与密码学》77(2)，第 301–319 页，2015 年。[11] H. D. Hollmann、J. H. Van Lint、J. P. Linnartz 和 L. M. Tolhuizen，《关于具有可识别父代性质的码》，《组合理论杂志 A辑》82(2)，第 121–133 页，1998 年。[12] W. Kautz 和 R. Singleton，《非随机二元叠加码》，IEEE 信息论汇刊 10，第 363–377 页，1964 年。[13] R.A. Moser 和 G. Tardos，《一般 Lov´asz 局部引理的构造性证明》，《ACM杂志》57，第 1–15 页，2010 年。[14] C. Shangguan 和 G. Ge，《可分离矩阵测试次数的新界》，IEEE 信息论汇刊 61(12)，第 7518–7521 页，2016 年。[15] C. Shangguan、X. Wang、G. Ge 和 Y. Miao，《抗诬陷码的新界》，IEEE 信息论汇刊 63，7252，2017 7247页。– ,[16] J. N. Staddon、D. R. Stinson 和 R. Wei，《抗诬陷码与可追踪码的组合性质》，IEEE 信息论汇刊 47(3)，第 1042–1049 页，2001 年。[17] D. R. Stinson 和 R. Wei，《可追踪方案与抗诬陷码的组合性质和构造》，SIAM 离散数学杂志 11(1)，第 41–53 页，1998 年。[18] D. R. Stinson、R. Wei 和 K. Chen，《关于广义分离哈希族》，《组合理论杂志 A辑》115，120，2008 年 105页。- .[19] D. R. Stinson 和 G. M. Zaverucha，《安全抗诬陷码及相关分离哈希族的一些改进界》，IEEE 信息论汇刊 54(6)，第 2508–2514 页，2008 年。[20] T. Van Trung，《从分离哈希族角度看抗诬陷码的紧界》，《设计、编码与密码学》72(3)，第 713–718 页，2014 年。