ELSEVIER

Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®

Discrete Mathematics 279 (2004) 463–477

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Generalized cover-free families

D.R. Stinson[a], R. Wei[b]

[a] *School of Computer Science, University of Waterloo, Waterloo, Ont., Canada N2L 3G1*
[b] *Department of Computer Science, Lakehead University, Thunder Bay, Ont., Canada P7B 5E1*

Received 15 November 2002; received in revised form 5 April 2003; accepted 9 June 2003

Abstract

Cover-free families have been investigated by many researchers, and several variations of these set systems have been used in diverse applications. In this paper, we introduce a generalization of cover-free families which includes as special cases all of the previously used definitions. Then we give several bounds and some efficient constructions for these generalized cover-free families.

Keywords: Cover-free family; Probabilistic method

1. Introduction

Cover-free families were first introduced in 1964 by Kautz and Singleton [13] in the context of superimposed binary codes. These structures have been discussed in several equivalent formulations, in subjects such as information theory, combinatorics and group testing, by numerous researchers (see, for example, [1,4,7–10,12,18,24,25]). Recently, cover-free families have been used to solve some new problems in cryptography and communications, including blacklisting, broadcast encryption, broadcast anti-jamming, source authentication in a network setting, and group key predistribution (see [2,3,6,11,15,17,19,20–23]). The original definition of cover-free families, as given in [13,9,12], was generalized in various ways, for example, in [5,17].

In this paper, we give a more general definition of cover-free families. We then investigate properties, bounds and constructions of these generalized cover-free families. All the previous definitions of cover-free families are special cases of the new definition.

In the rest of this section, we give the definitions and notations used in this paper. In Section 2, we discuss bounds (i.e., necessary conditions) for generalized cover-free families, which are obtained through two different approaches. In Section 3, we give

E-mail addresses: dstinson@uwaterloo.ca (D.R. Stinson), wei@ccc.cs.lakeheadu.ca (R. Wei).

0012-365X/$ - see front matter
doi:10.1016/S0012-365X(03)00287-5

some good explicit constructions of generalized cover-free families, as well as non-constructive existence results using the probabilistic method.

1.1. Definitions and notations

A *set system* is a pair $(X,\mathscr{F})$, where X is a set of *points* and $\mathscr{F}$ is a set of subsets of X (called *blocks*). A set system $(X,\mathscr{F})$ is called *k-uniform* if $|F|=k$ for each $F\in\mathscr{F}$. Throughout this paper, we will use N and T to denote the cardinality of X and $\mathscr{F}$ respectively.

Now we give a general definition of cover-free families, as follows.

Definition 1.1. Let w, r and d be positive integers. A set system $(X,\mathscr{F})$ is called a $(w,r;d)$-*cover-free family* (or $(w,r;d)$-*CFF*) provided that, for any w blocks $B_1,\ldots,B_w\in\mathscr{F}$ and any other r blocks $A_1,\ldots,A_r\in\mathscr{F}$, we have that

$$\left|\left(\bigcap_{i=1}^{w}B_i\right)\setminus\left(\bigcup_{j=1}^{r}A_j\right)\right|\geqslant d.$$

Less formally, the intersection of any w blocks contains at least d points that are not in the union of r other blocks.

Sometimes, we will use the notation $(w,r;d)$-CFF(N,T) to denote a cover-free family in which $|X|=N$ and $|\mathscr{F}|=T$ (i.e., there are N points and T blocks). $(1,r;1)$-CFF were defined in [13,9,12] for different purposes. $(w,r;1)$-CFF were defined in [17] for some cryptographic applications (namely, to permit the construction of certain key distribution schemes). $(1,r;d)$-CFF were defined in [5] in connection with superimposed distance codes. $(w,r;d)$-CFF for general w,r and d were first considered in [20]; however, the equivalent dual version of cover-free families (disjunct families; see below) was used in that paper, and $(w,r;d)$-CFF were not explicitly defined there.

A set system can be described by an incidence matrix. Let $(X,\mathscr{B})$ be a set system where $X=\{x_1,x_2,\ldots,x_N\}$ and $\mathscr{B}=\{B_1,B_2,\ldots,B_T\}$. The *incidence matrix* of $(X,\mathscr{B})$ is the $N\times T$ matrix $A=(a_{ij})$, where

$$a_{ij}=\begin{cases}1 & \text{if } x_i\in B_j,\\ 0 & \text{if } x_i\notin B_j.\end{cases}$$

Conversely, given an incidence matrix, we can define an associated set system in an obvious way.

Disjunct systems and cover-free families are dual incidence structures. If A is an incidence matrix of a cover-free family, then A^{T}, the transpose of A, is an incidence matrix of a disjunct system. We have the following definition of (generalized) disjunct systems.

Definition 1.2. A set system $(X,\mathscr{B})$ is a $(w,r;d)$-*disjunct system* provided that, for any $P,Q\subseteq X$ such that $|P|\leqslant w$, $|Q|\leqslant r$ and $P\cap Q=\emptyset$, there exist at least d blocks $B\in\mathscr{B}$

such that $P \subseteq B$ and $Q \cap B = \emptyset$. A $(w,r;d)$-disjunct system, $(X,\mathscr{B})$, will be denoted as a $(w,r;d)$-DS(N,T) if $|X| = N$ and $|\mathscr{B}| = T$.

From the above discussion, we have the following theorem.

Theorem 1.1. *There exists a* $(w,r;d)$-CFF(N,T) *if and only if there exists a* $(w,r;d)$-DS(T,N).

2. Bounds for cover-free families

It is easy to see that there is a tradeoff between the values of N and T in a cover-free family. We want to maximize the value of T or minimize the value of N in the case that the other parameters are given. For given values w,r,d and T, let $N((w,r;d),T)$ denote the minimum value of N such that a $(w,r;d)$-CFF(N,T) exists. Similarly, let $T((w,r;d),N)$ denote the maximum value of T such that a $(w,r;d)$-CFF(N,T) exists.

2.1. Bounds from coverings of hypergraphs

Engel proved some bounds for $(w,r;1)$-CFF in [7]. Now we generalize these results to $(w,r;d)$-CFF for general d.

Denote $[n]=\{1,\ldots,n\}$. Define $P_{n;l,u}=\{X \subseteq [n] : l \leqslant |X| \leqslant u\}$, where $0 < l < u < n$, in which the sets in $P_{n;l,u}$ are ordered by inclusion. Define a class of *order-interval hypergraphs* $G_{n;l,u} = (P,E)$ as follows. Let the set of points be $P = P_{n;l,u}$, and let the set of edges E be the *maximal intervals*, i.e., for any $X,Y \subseteq [n]$ such that $|X| = l$ and $|Y| = u$, define

$$I(X,Y) = \{C \subseteq [n] : X \subseteq C \subseteq Y\},$$

and then define

$$E = \{I(X,Y) : |X| = l, |Y| = u, X, Y \subseteq [n]\}.$$

Definition 2.1. A *point d-cover*, or simply a *d-cover*, of a hypergraph is a subset of points $\mathscr{C}$ such that each edge of the hypergraph contains at least d points of $\mathscr{C}$.

The following theorem shows the equivalence of CFF and a certain covering of a $G_{n,l,u}$. The result is phrased in terms of disjunct systems.

Theorem 2.1. *There exists a d-cover of* $G_{n,l,u}$ *of size b if and only if there exists an* $(l, n-u; d)$-DS(n,b).

Proof. $\mathscr{C}$ is a d-cover of $G_{n,l,u}$ if and only if, for any $X,Y \subseteq [n]$ such that $|X| = l$ and $|Y| = u$, there are d points $C \in \mathscr{C}$ such that $X \subseteq C \subseteq Y$. Equivalently, for any $X,Z \subseteq [n]$ such that $|X| = l$ and $|Z| = n-u$, there are d points $C \in \mathscr{C}$ such that $X \subseteq C$ and $Z \cap C = \emptyset$. This is the same thing as a $(l, n-u; d)$-DS(n,b). □

Let $\tau^d_{n;l,u} = \min\{|\mathscr{C}| : \mathscr{C} \text{ is a point } d\text{-cover of } G_{n;l,u}\}$. Then we have the following corollary.

Corollary 2.2. $N((w,r;d),T) = \tau^d_{T;w,T-r}$.

We need some tools from graph theory. A *fractional d-cover* is a function $g : P \to \mathbb{R}^+$, such that, for any $I(X,Y) \in E$, it holds that

$$\sum_{Z \in I(X<Y)} g(Z) \geqslant d,$$

where $\mathbb{R}^+$ is the set of nonnegative real numbers.

Define the *fractional d-covering number* $(\tau^*)^d_{n;l,u}$ as follows:

$$(\tau^*)^d_{n;l,u} = \min\left\{ \sum_{Z \in P} g(Z) : g \text{ is a fractional } d\text{-cover of } G_{n;l,u} \right\}.$$

When $d = 1$, we write $(\tau^*)_{n;l,u}$ instead of $(\tau^*)^1_{n;l,u}$. The following lemma gives the relationship of $\tau^d_{n;l,u}$ and $(\tau^*)^d_{n;l,u}$.

Lemma 2.3. $\tau^d_{n;l,u} \geqslant (\tau^*)^d_{n;l,u}$.

Proof. Suppose $\mathscr{C}$ is a d-cover of $G_{n;l,u}$. Define

$$g(Z) = \begin{cases} 1 & \text{if } Z \in \mathscr{C}, \\ 0 & \text{otherwise.} \end{cases}$$

Then g is a fractional d-cover. Therefore $|\mathscr{C}| \geqslant (\tau^*)^d_{n;l,u}$. □

The paper [7] only considered the case $d = 1$. For $d > 1$, we can use the following lemma.

Lemma 2.4. $(\tau^*)^d_{n;l,u} = d \times (\tau^*)_{n;l,u}$.

Proof. Suppose that g is a fractional 1-cover of size $(\tau^*)_{n;l,u}$. Define

$$g'(Z) = d \times g(Z)$$

for all $Z \in P$. Then g' is a fractional d-cover. Hence we have

$$(\tau^*)^d_{n;l,u} \leqslant d \times (\tau^*)_{n;l,u}.$$

In a similar way, we can prove that

$$(\tau^*)_{n;l,u} \leqslant \frac{1}{d} \times (\tau^*)^d_{n;l,u}. \qquad \square$$

Now we can give a formula for the fractional-covering numbers, as follows:

Lemma 2.5.

$$(\tau^*)^d_{n;l,u} = d \times \min\left\{ \binom{n}{m} \Big/ \binom{u-l}{m-l} : l \leqslant m \leqslant u \right\}.$$

Proof. It is proved that

$$(\tau^*)_{n;l,u} = \min\left\{ \binom{n}{m} \Big/ \binom{u-l}{m-l} : l \leqslant m \leqslant u \right\}$$

in [7]. The conclusion follows from Lemma 2.4. □

Therefore, from Lemmas 2.3 and 2.5, we have the following bound for cover-free families.

Corollary 2.6.

$$N((w,r;d),T) \geqslant \min\left\{ d \binom{T}{m} \Big/ \binom{T-r-w}{m-w} : w \leqslant m \leqslant T-r \right\}.$$

The following two theorems are generalizations of [7, Proposition 3].

Theorem 2.7. *For $\lambda < l$ and $\mu > u$, it holds that $\tau^d_{n;l,u} \geqslant (\tau^*)^d_{n;\lambda,\mu} \times (\tau^*)_{\mu-\lambda;l-\lambda,u-\lambda}$.*

Proof. Let $\mathscr{C}$ be an optimal d-cover of $G_{n;l,u}$. Since $\lambda < l$ and $\mu > u$, $\mathscr{C}$ is also a set of points in $G_{n;\lambda,\mu}$. In $G_{n;\lambda,\mu}$ define

$$g(Z) = \begin{cases} \dfrac{1}{(\tau^*)_{\mu-\lambda;l-\lambda,u-\lambda}} & \text{if } Z \in \mathscr{C}, \\ 0 & \text{otherwise.} \end{cases}$$

We are going to prove that g is a fractional d-cover of $G_{n;\lambda,\mu}$. Suppose $I(Y_1,Y_2)$ is an edge of $G_{n;\lambda,\mu}$; then $|Y_1| = \lambda$ and $|Y_2| = \mu$. So

$$I(Y_1,Y_2) \cap P_{n;l,u} = \{Z : Y_1 \subseteq Z \subseteq Y_2, l \leqslant |Z| \leqslant u\},$$

which is isomorphic to $P_{\mu-\lambda;l-\lambda,u-\lambda}$. It is obvious that $\mathscr{C} \cap I(Y_1,Y_2)$ gives rise to a d-cover of $G_{\mu-\lambda;l-\lambda,u-\lambda}$ by deleting the elements of Y_1 from every point $C \in \mathscr{C} \cap I(Y_1,Y_2)$. Therefore, we have

$$\sum_{Z \in I(Y_1,Y_2)} g(Z) = \frac{|\mathscr{C} \cap I(Y_1,Y_2)|}{(\tau^*)_{\mu-\lambda;l-\lambda,u-\lambda}} \geqslant \frac{(\tau^*)^d_{\mu-\lambda;l-\lambda,u-\lambda}}{(\tau^*)_{\mu-\lambda;l-\lambda,u-\lambda}} = d,$$

where the last equality comes from Lemma 2.4. Now, since

$$(\tau^*)^d_{n;\lambda,\mu} \leqslant \sum_{Z \in P_{n;\lambda,\mu}} g(Z) = \sum_{Z \in P_{n;l,u}} g(Z) = \frac{\tau^d_{n;l,u}}{(\tau^*)_{\mu-\lambda;l-\lambda,u-\lambda}},$$

the conclusion follows. □

We now prove a bound for cover-free families using Theorem 2.7.

Theorem 2.8. *For $0 < \lambda_1 < w, 0 < \lambda_2 < r$, we have that*

$$N((w,r;d),T) \geqslant d\frac{\binom{T}{m_1}\binom{T-w-r+\lambda_1+\lambda_2}{m_2}}{\binom{T-w-r+\lambda_1+\lambda_2}{m_1-w+\lambda_1}\binom{T-w-r}{m_2-\lambda_1}},$$

where m_1 and m_2 are chosen such that the right side of the above inequality attains its minimum value, subject to the constraints that $w-\lambda_1 \leqslant m_1 \leqslant T-r+\lambda_2$, $\lambda_1 \leqslant m_2 \leqslant T-w-r+\lambda_1$, and m_1 and m_2 are integers.

Proof. Apply Theorem 2.7, letting $n=T$, $l=w$, $u=T-r$, $\lambda=w-\lambda_1$ and $\mu=T-r+\lambda_2$. □

The following theorem gives a recursive method to bound the value of τ.

Theorem 2.9. *For $\lambda < l$ and $\mu > u$, it holds that $\tau^d_{n;l,u} \geqslant (\tau^*)_{n;\lambda,\mu} \times \tau^d_{\mu-\lambda;l-\lambda,u-\lambda}$.*

Proof. The proof is similar to that of Theorem 2.7. In this proof, we define

$$g(Z) = \begin{cases} \dfrac{1}{\tau^d_{\mu-\lambda;l-\lambda,u-\lambda}} & \text{if } Z \in \mathscr{C}, \\ 0 & \text{otherwise.} \end{cases}$$

Then it can be shown that g is a fractional 1-cover of $G_{n;\lambda,\mu}$. □

As a corollary of the above theorem, we obtain the following result by applying Lemma 2.5.

Theorem 2.10. *For $0 < \lambda_1 < w, 0 < \lambda_2 < r$, it holds that*

$$N((w,r;d),T) \geqslant \frac{\binom{T}{m}}{\binom{T-w-r+\lambda_1+\lambda_2}{m-w+\lambda_1}} N((\lambda_1,\lambda_2;d),T-w-r+\lambda_1+\lambda_2),$$

where m is chosen such that the right side of the above inequality attains its minimum value, subject to the constraints that $w-\lambda_1 \leqslant m \leqslant T-r+\lambda_2$ and m is an integer.

Now set $\lambda_1 = w-1$ and $\lambda_2 = r-1$ in the above theorem. Then it is easy to check that the right side of the inequality attains its minimum value when $m = T/2$. So we have the following result.

Lemma 2.11.

$$N((w,r;d),T) \geqslant 4\left(1-\frac{1}{T}\right)N((w-1,r-1;d),T-2).$$

In [5,26], the following bound for $(1,r;d)$-CFF was proven:

Theorem 2.12. *If $r > 1$ and $d \geqslant 1$ are integers, then*

$$N((1,r;d),T) \geqslant c\left(\frac{r^2}{\log r}\log T + (d-1)r\right),$$

where c is some constant.

Combining Theorem 2.12 and Lemma 2.11, we are able to prove the following bound.

Theorem 2.13. *Suppose r, w and d are integers, $r > w \geqslant 1$, and $d \geqslant 1$. Then*

$$N((w,r;d),T) \geqslant c4^{w-1}\left(1-\frac{1}{T}\right)\left(1-\frac{1}{T-2}\right)\cdots\left(1-\frac{1}{T-2w+2}\right) \times\left(\frac{(r-w+1)^2}{\log(r-w+1)}\log(T-2w)+(d-1)(r-w+1)\right)$$

for some constant c.

Proof. Iterate Lemma 2.11 $w-1$ times, and then apply Theorem 2.12. □

2.2. Bound from recursive methods

Recently, two bounds for $(w,r;1)$-CFF were proven in [24]. We use the same techniques to give bounds on $(w,r;d)$-CFF, for general d. The following two simple lemmas can be proven in a similar way as the corresponding lemmas in [24].

Lemma 2.14.

$$N((w,r;d),T) \geqslant N((w,r-1;d),T-1) + N((w-1,r;d),T-1).$$

Lemma 2.15.

$$N((w,r;d),T) = N((r,w;d),T).$$

To discuss the first bound, we define

$$g(w,r,T) = \frac{\binom{w+r}{w}\log T}{\log(w+r)}.$$

The function g satisfies the following property (see [24]):

Lemma 2.16. *For $w, r \geqslant 2$ and $T \geqslant w+r$, it holds that*

$$g(w,r,T) \leqslant g(w,r-1,T-1) + g(w-1,r,T-1).$$

We need a simple numerical lemma, which was also proven in [24].

Lemma 2.17. *For $r \geqslant 2$, it holds that*

$$\frac{r^2}{\log r} \geqslant \frac{2r+2}{\log(r+1)}.$$

Our bound is as follows.

Theorem 2.18. *For $w, r \geqslant 1$ and $T \geqslant w + r > 2$, it holds that*

$$N((w,r;d),T) \geqslant 2c \frac{\binom{w+r}{w}}{\log(w+r)} \log T + \frac{1}{2} c \binom{w+r}{w} (d-1),$$

where c is the same constant as in Theorem 2.12.

Proof. First, consider the case $w = 1$. From Theorem 2.12, we have

$$N((w,r;d),T) \geqslant c \frac{r^2}{\log r} \log T + \frac{1}{2} c(r+1)(d-1).$$

In this case, the conclusion follows from Lemma 2.17. Also, the case $r = 1$ is similar, in view of Lemma 2.15.

For the general case, where $r, w \geqslant 2$, we prove the bound by induction on $w + r$, as follows:

$$\begin{aligned} N((w,r;d),T) &\geqslant N((w-1,r;d),T-1) + N((w,r-1;d),T-1) \\ &\geqslant 2cg(w-1,r,T-1) + \frac{1}{2} c \binom{w+r-1}{w-1} (d-1) \\ &\quad + 2cg(w,r-1,T-1) + \frac{1}{2} c \binom{w+r-1}{r-1} (d-1) \\ &\geqslant 2cg(w,r,T) + \frac{1}{2} c \binom{w+r}{w} (d-1). \end{aligned}$$

Here, the first inequality comes from Lemma 2.14, the second one comes from an induction assumption, and the third one comes from Lemma 2.16. □

Another bound for $(w,r;1)$-CFF from [24] can also be generalized to $d > 1$. The proof is similar to that of [24, Theorem 4.4], and we omit the details here.

Theorem 2.19. *For any integers $w, r \geqslant 1$, there exists an integer $T_{w,r}$ such that*

$$N((w,r;d),T) \geqslant 0.7c \frac{\binom{w+r}{w}(w+r)}{\log\binom{w+r}{w}} \log T + \frac{1}{2} c \binom{w+r}{w} (d-1)$$

for all $T > T_{w,r}$, where c is the same constant as in Theorem 2.12.

Finally, we note that explicit upper bounds on the above-mentioned constants $T_{w,r}$ are given by Ma [16].

3. Constructions of cover-free families

It is easy to construct a $(w,r;d)$-CFF(dN,T) from a $(w,r;1)$-CFF(N,T). In fact, if $(X,\mathscr{F})$ is a $(w,r;1)$-CFF(N,T), then we can construct a $(w,r;d)$-CFF(dN,T) on $X\times\{0,1,\ldots,d-1\}$ by taking d copies of every point in every block. In this section, we discuss several more efficient constructions for $(w,r;d)$-CFF(N,T).

3.1. Combinatorial constructions

Several types of codes and combinatorial designs have been used to construct CFF. For example, t-designs were used in [8,13,23] to construct $(1,r;1)$-CFF. Orthogonal arrays have also been used, in [23,26], to construct CFF. The following result is obtained from orthogonal arrays (see [26] for the details).

Theorem 3.1. *For any prime power q, and any integer t such that $2\leqslant t<q$, there exists a $(1,\lfloor(q-d)/(t-1)\rfloor;d)$-CFF$(q^2+q,q^t)$.*

The following construction is a generalization of Sperner's Theorem.

Lemma 3.2.

$$T((1,1;d)N)\geqslant\binom{\lfloor\frac{N}{d}\rfloor}{\lfloor\frac{N}{2d}\rfloor}.$$

Proof. Let $S_1,S_2,\ldots,S_{\lfloor N/d\rfloor}$ be a partition of the set $\{0,1,\ldots,N-1\}$, such that $|S_i|=d$ for $i=1,\ldots,\lfloor N/d\rfloor$. Define blocks as follows:

$$F_{i_1,\ldots,i_k}=\bigcup_{j=1}^{k}S_{i_j},$$

for all k-subsets of $\{1,2,\ldots,\lfloor N/d\rfloor\}$. Then it is easy to check that the resulting system is a $(1,1;d)$-CFF. Therefore it holds that

$$T((1,1;d)N)\geqslant\binom{\lfloor\frac{N}{d}\rfloor}{k}$$

and the conclusion follows. □

When $d=1$, the above construction is optimal. In fact, it is just a Sperner system.

In [22], separating hash families are used to construct $(w,r;1)$-CFF. We now generalize this method to construct $(w,r;d)$-CFF.

Definition 3.1. An $(n,m,\{w_1,w_2\})$-*d-separating hash family* is a set of functions $\mathscr{F}$, such that $|Y|=n$, $|X|=m$, $f: Y \to X$ for each $f \in \mathscr{F}$, and for any $C_1, C_2 \subseteq \{1,2,\ldots,n\}$ such that $|C_1| = w_1$, $|C_2| = w_2$ and $C_1 \cap C_2 = \emptyset$, there exist at least d functions $f \in \mathscr{F}$ such that

$$\{f(y): y \in C_1\} \cap \{f(y): y \in C_2\} = \emptyset.$$

The notation d-SHF$(N;n,m,\{w_1,w_2\})$ will be used to denote an $(n,m,\{w_1,w_2\})$-d-separating hash family with $|\mathscr{F}| = N$.

A d-SHF$(N;n,m,\{w_1,w_2\})$ can be depicted as an $N \times n$ matrix with entries from $\{1,2,\ldots,m\}$, such that in any two disjoint sets C_1 and C_2 of w_1 and w_2 columns (respectively), there exist at least d rows such that the entries in the columns C_1 are distinct from the entries in the columns C_2.

Now suppose that A is the $n \times N$ transposed matrix which is derived from a d_1-SHF$(N;n,m,\{w,r\})$. The elements in A are denoted $1,2,\ldots,m$. Suppose that B is the incidence matrix of a $(w,r;d_2)$-CFF(v,m). Denote the rows of B by $b_1,b_2,\ldots,b_m$. We construct an $Nv \times n$ matrix A' by replacing every element i in the array A by the row vector b_i. It can be verified that A' is the incidence matrix of a $(w,r;d_1d_2)$-CFF(vN,n). Thus we have the following.

Theorem 3.3. *If there exists a* $(w,r;d_1)$-CFF(v,m) *and a* d_2-SHF$(N;n,m,\{w,r\})$, *then there exists a* $(w,r;d_1d_2)$-CFF(vN,n).

The function $\log^*$ is defined recursively as follows:

$$\log^*(1) = 1,$$

$$\log^*(n) = \log^*(\lceil \log n \rceil) + 1 \quad \text{if } n > 1.$$

Using orthogonal arrays that are easily constructed from Reed-Solomon codes, and a recursive method based on Theorem 3.3, the following result is proven in [24].

Theorem 3.4. *Let* m, w_1 *and* w_2 *be positive integers. Then there exists an infinite class of* 1-SHF$(N;n,m,\{w_1,w_2\})$, *for which* N *is* $O((w_1w_2)^{\log^*(n)}(\log n))$.

Note that, for any w,r and d, we can construct a $w+r$ by $d\binom{w+r}{w}$ matrix by taking d copies of every possible $0-1$ column vector having hamming weight equal to w. Then we obtain a $(w,r;d)$-CFF$\left(d\binom{w+r}{w}, w+r\right)$. From Theorems 3.4 and 3.3, we obtain the following result.

Theorem 3.5. *For any positive integers* w,r *and* d, *there exists a* $(w,r;d)$-CFF $\left(d\binom{w+r}{w}N,T\right)$, *where* N *is* $O((w_1w_2)^{\log^*(T)}(\log T))$.

Another well-known combinatorial object known as a perfect hash family can also be used to construct CFF.

Definition 3.2. An (n,m,w)-*d-perfect hash family* is a set of functions $\mathscr{F}$, such that $|Y|=n$, $|X|=m$, $f:Y\to X$ for each $f\in\mathscr{F}$, and for any $C\subseteq\{1,2,\ldots,n\}$ such that $|C|=w$, there exist at least d functions $f\in\mathscr{F}$ such that f is one-to-one on C. The notation d-PHF$(N;n,m,w)$ will be used to denote an (n,m,w)-d-perfect hash family with $|\mathscr{F}|=N$.

Since any d-PHF$(N;n,m,w+r)$ is automatically a d-SHF$(N;n,m,\{w,r\})$, we can apply Theorem 3.3 using PHF as ingredients. There are many papers providing explicit constructions of PHF. For example, we can use a construction of PHF from algebraic curves over finite fields, which was described in [25].

Theorem 3.6. *For any positive integers $m\geqslant w$, there exists an explicit construction for an infinite family of* 1-PHF$(N;n,m,w)$ *such that N is $O(\log n)$.*

From Theorems 3.6 and 3.3, we have the following result.

Theorem 3.7. *For any positive integers w,r and d, there exists an explicit construction for an infinite family of* $(w,r;d)$-CFF$\left(d\binom{w+r}{w}N,T\right)$, *where N is $O(\log T)$.*

3.2. Non-constructive existence results

Probabilistic methods have been used by many researchers to establish the existence of "good" cover-free families. Here we consider the existence of generalized CFF.

We will construct an $N\times T$ matrix which satisfies the conditions of Definition 1.2. Let A be an $N\times T$ $0-1$ matrix whose columns are labelled $1,\ldots,T$. Suppose that $C_1,C_2\subseteq\{1,\ldots,T\}$, $|C_1|=w$, $|C_2|=r$ and $C_1\cap C_2=\emptyset$. Define $\mathbf{X}_A(C_1,C_2)=0$ if there exist at least d rows of A such that the entries in the columns in C_1 are all "1"s and the entries in the columns in C_2 are all "0"s, and define $\mathbf{X}_A(C_1,C_2)=1$, otherwise.

Suppose A is an $N\times T$ matrix in which each entry is defined to be a "1" with probability ρ. (The value of ρ will be chosen a bit later.) We say that a row is "good" if the entries in the columns in C_1 are all "1" and the entries in the columns in C_2 are all "0". The probability that a particular row is good is

$$p=\rho^w(1-\rho)^r.$$

To maximize the value of p, we let

$$\rho=\frac{w}{w+r}.$$

We will make use of the following "tail inequality" (see [14, p. 106]) which can be seen as a special case of the Chernoff Bound.

Lemma 3.8. *Suppose $\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_n$ are independent random variables such that $Pr[\mathbf{X}_i=1]=p$ and $Pr[\mathbf{X}_i=0]=1-p$ for $1\leqslant i\leqslant n$. Let $\mathbf{X}=\mathbf{X}_1+\mathbf{X}_2+\cdots+\mathbf{X}_n$. Then for any $a\geqslant 0$, it holds that*

$$Pr[\mathbf{X}\leqslant n(p-a)]\leqslant \mathrm{e}^{-a^2n/2p}.$$

In order to apply Lemma 3.8, define

$$\mathbf{X}_i = \begin{cases} 1 & \text{if the } i\text{th row of } A \text{ is good,} \\ 0 & \text{otherwise,} \end{cases}$$

for $1 \leqslant i \leqslant N$, and define $n = N$. Then we have that

$$\mathsf{Exp}[\mathbf{X}_A(C_1, C_2)] = Pr[\mathbf{X} \leqslant d - 1].$$

Let $a = p/2$ and $N = 2(d-1)/p$. Then $N(p-a) = d-1$, and applying Lemma 3.8, we have that

$$\mathsf{Exp}[\mathbf{X}_A(C_1, C_2)] \leqslant \mathrm{e}^{-pN/8}.$$

Now, if we define the random variable

$$\mathbf{X}_A = \sum_{\{C_1, C_2 \subseteq \{1,\ldots,T\} : |C_1|=w, |C_2|=r, C_1 \cap C_2 = \emptyset\}} \mathbf{X}_A(C_1, C_2),$$

then it is easy to see that

$$\begin{aligned} \mathsf{Exp}[\mathbf{X}_A] &\leqslant \binom{T}{w}\binom{T-w}{r}\mathrm{e}^{-pN/8} \\ &< \frac{T^{w+r}}{w!r!}\mathrm{e}^{-pN/8}. \end{aligned}$$

If $\mathsf{Exp}[\mathbf{X}_A] < 1$, then a CFF exists. Therefore, we obtain the following theorem about the existence of generalized CFF.

Theorem 3.9. *Suppose that T, w and r are positive integers. Define*

$$p = \frac{w^w r^r}{(r+w)^{r+w}}.$$

If

$$N > \frac{8}{p}((w+r)\log T - \log w! - \log r!)$$

then there exists a $(w, r; d)$-CFF(N, T) for

$$d = \frac{pN}{2} + 1.$$

Proof. Suppose

$$N > \frac{8}{p}((w+r)\log T - \log w! - \log r!).$$

Then

$$\frac{T^{w+r}}{w!r!}\,\mathrm{e}^{-pN/8} < 1,$$

which implies that $\mathsf{Exp}[\mathbf{X}_A] < 1$, and hence the desired CFF exists. □

By instead considering the inequality

$$\frac{T^{w+r}}{w!r!}\,\mathrm{e}^{-pN/8} < \mathrm{e}^{-t}$$

in the above proof, we obtain the following result.

Theorem 3.10. *Suppose that T, w and r are positive integers. Define*

$$p = \frac{w^w r^r}{(r+w)^{r+w}}.$$

If

$$N > \frac{8}{p}((w+r)\log T + t - \log w! - \log r!),$$

then the probability that the matrix A is not a $(w,r;d)$-CFF(N,T), for

$$d = \frac{pN}{2} + 1,$$

is at most e^{-t}.

Next, we use the probabilistic method to prove an existence result for k-uniform $(r,w;d)$-CFF. We adapt a similar method which was used in [15]. Fix an integer $\ell \geqslant 2$, and let $N = k\ell$. As before, let A be an $N \times T$ $0-1$ matrix whose columns are labelled $1,\ldots,T$. Suppose that $C_1, C_2 \subseteq \{1,\ldots,T\}$, $|C_1| = w$, $|C_2| = r$ and $C_1 \cap C_2 = \emptyset$. Also, define $\mathbf{X}_A(C_1,C_2)$ as before.

Suppose that A is partitioned into k disjoint $\ell \times T$ subarrays which are denoted A_i, $1 \leqslant i \leqslant k$. Each column of each A_i is chosen to be a random $0-1$ column vector of length ℓ having hamming weight equal to 1. We say that a subarray A_i is "good" if there exists a row of A_i such that the entries in the columns in C_1 are all "1" and the entries in the columns in C_2 are all "0". (Notice that every A_i can contain at most one such row.) The probability that a particular A_i is good is

$$p = \left(\frac{1}{\ell}\right)^{w-1}\left(1-\left(\frac{1}{\ell}\right)\right)^r = \frac{(\ell-1)^r}{\ell^{w+r-1}}.$$

Define

$$\mathbf{X}_i = \begin{cases} 1 & \text{if } A_i \text{ is good,} \\ 0 & \text{otherwise,} \end{cases}$$

for $1 \leqslant i \leqslant k$, and define $n = k$ (i.e., $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_k$). Then, we have that

$$\mathsf{Exp}[\mathbf{X}_A(C_1,C_2)] = Pr[\mathbf{X} \leqslant d-1].$$

Let $a = p/2$ and $k = 2(d-1)/p$. Applying Lemma 3.8, we have that

$$\mathsf{Exp}[\mathbf{X}_A(C_1, C_2)] \leqslant \mathrm{e}^{-pk/8}.$$

Now, if we define the random variable

$$\mathbf{X}_A = \sum_{\{C_1,C_2\subseteq\{1,\ldots,T\}\,:\,|C_1|=w,|C_2|=r,C_1\cap C_2=\emptyset\}} \mathbf{X}_A(C_1, C_2),$$

then it is easy to see that

$$\mathsf{Exp}[\mathbf{X}_A] \leqslant \binom{T}{w}\binom{T-w}{r}\mathrm{e}^{-pk/8} < \frac{T^{w+r}}{w!r!}\mathrm{e}^{-pk/8}.$$

We obtain the following theorem about the existence of k-uniform generalized CFF.

Theorem 3.11. *Suppose that T, w, r and ℓ are positive integers. Define*

$$p = \frac{(\ell-1)^r}{\ell^{w+r-1}}.$$

If

$$k > \frac{8}{p}((w+r)\log T - \log w! - \log r!)$$

then there exists a k-uniform $(w, r; d)$-CFF$(k\ell, T)$ for

$$d = \frac{pk}{2} + 1.$$

The following variation is proved in a similar fashion.

Theorem 3.12. *Suppose that T, w, r and ℓ are positive integers. Define*

$$p = \frac{(\ell-1)^r}{\ell^{w+r-1}}.$$

If

$$k > \frac{8}{p}((w+r)\log T + t - \log w! - \log r!)$$

then the probability that the matrix A is not a k-uniform $(w, r; d)$-CFF$(k\ell, T)$, for

$$d = \frac{pk}{2} + 1,$$

is at most e^{-t}.

Acknowledgements

Tha authors' research is supported as follows: NSERC grants RGPIN 203114-02 and IRC 216431-96 (DRS); and NSERC grant RGPIN 239135-01 (RW).

References

[1] K.A. Bush, W.T. Federer, H. Pesotan, D. Raghavarao, New combinatorial designs and their application to group testing, J. Statist. Plann. Inference 10 (1984) 335–343.

[2] R. Canetti, J. Garay, G. Itkis, D. Micciancio, M. Naor, B. Pinkas, Multicast security: a taxonomy and some efficient constructions, In Proceedings of INFOCOM '99, Vol. 2, pp. 708–716.

[3] Y. Desmedt, R. Safavi-Naini, H. Wang, L. Batten, C. Charnes, J. Pieprzyk, Broadcast anti-jamming systems, Comput. Networks 35 (2001) 223–236.

[4] A.G. Dyachkov, V.V. Rykov, Bounds on the length of disjunctive codes, Problemy Peredachi Informatsii 18 (1982) 7–13 (in Russian).

[5] A.G. Dyachkov, V.V. Rykov, A.M. Rashad, Superimposed distance codes, Problems Control Inform. Theory 18 (1989) 237–250.

[6] M. Dyer, T. Fenner, A. Frieze, A. Thomason, On key storage in secure networks, J. Cryptol. 8 (1995) 189–200.

[7] K. Engel, Interval packing and covering in the boolean lattice, Combin. Probab. Comput. 5 (1996) 373–384.

[8] P. Erdös, P. Frankl, Z. Füredi, Families of finite sets in which no set is covered by the union of two others, J. Combin. Theory Ser. A 33 (1982) 158–166.

[9] P. Erdös, P. Frankl, Z. Füredi, Families of finite sets in which no set is covered by the union of r others, Israel J. Math. 51 (1985) 75–89.

[10] Z. Füredi, On r-cover-free families, J. Combin. Theory Ser. A 73 (1996) 172–173.

[11] J.A. Garay, J. Staddon, A. Wool, Long-lived broadcast encryption, in Advances in Cryptology–Crypto 2000, Lecture Notes in Computer Science, Vol. 1880, Springer, Berlin, 2000, pp. 333–352.

[12] F.K. Hwang, V.T. Sós, Non-adaptive hypergeometric group testing, Studia Sci. Math. Hungar. 22 (1987) 257–263.

[13] W.H. Kautz, R.C. Singleton, Nonrandom binary superimposed codes, IEEE Trans. Inform. Theory 10 (1964) 363–377.

[14] D.E. Knuth, The Art of Computer Programming, 3rd Edition, Addison Wesley, Reading, MA, 1997.

[15] R. Kumar, S. Rajagopalan, A. Sahai, Coding constructions for blacklisting problems without computational assumptions, In Advances in Cryptology–Crypto '99, Lecture Notes in Computer Science, Vol. 1666, Springer, Berlin, 1999, pp. 609–623.

[16] X. Ma, An evaluative problem related to a new bound of cover-free family, preprint.

[17] C.J. Mitchell, F.C. Piper, Key storage in secure networks, Discrete Appl. Math. 21 (1988) 215–228.

[18] M. Ruszinkó, On the upper bound of the size of the r-cover-free families, J. Combin. Theory Ser. A 66 (1994) 302–310.

[19] R. Safavi-Naini, H. Wang, Multireceiver authentication codes: models, bounds, constructions, and extensions, Inform. Comput. 151 (1999) 148–172.

[20] D.R. Stinson, On some methods for unconditionally secure key distribution and broadcast encryption, Des. Codes Cryptogr. 12 (1997) 215–243.

[21] D.R. Stinson, Tran van Trung, Some new results on key distribution patterns and broadcast encryption, Des. Codes Cryptogr. 14 (1998) 261–279.

[22] D.R. Stinson, Tran van Trung, R. Wei, Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, J. Statist. Plann. Inference 86 (2000) 595–617.

[23] D.R. Stinson, R. Wei, Combinatorial properties and constructions of traceability schemes and frameproof codes, SIAM J. Discrete Math. 11 (1998) 41–53.

[24] D.R. Stinson, R. Wei, L. Zhu, Some new bounds for cover-free families, J. Combin. Theory Ser. A 90 (2000) 224–234.

[25] H. Wang, C. Xing, Explicit constructions of perfect hash families from algebraic curves over finite fields, J. Combin. Theory Ser. A 93 (2001) 112–124.

[26] R. Wei, On cover-free families, preprint.