

Tracing a single user

Noga Alon^{*} Vera Asodi[†]

Abstract

Let $g(n, r)$ be the maximum possible cardinality of a family \mathcal{F} of subsets of $\{1, 2, \dots, n\}$ so that given a union of at most r members of \mathcal{F} , one can identify at least one of these members. The study of this function is motivated by questions in molecular biology. We show that $g(n, r) = 2^{\Theta(\frac{n}{r})}$, thus solving a problem of Csürös and Ruszinkó.

1 Introduction

Let $[n] = \{1, 2, \dots, n\}$, and let $\mathcal{F} \subseteq 2^{[n]}$ be a family of subsets of $[n]$. \mathcal{F} is called r -superimposed if for all $A_1, \dots, A_k, B_1, \dots, B_l \in \mathcal{F}$ with $k, l \leq r$ and $\{A_1, \dots, A_k\} \neq \{B_1, \dots, B_l\}$,

$$\bigcup_{i=1}^k A_i \neq \bigcup_{i=1}^l B_i.$$

This means that, given the union of up to r sets from an r -superimposed family, one can identify all those sets. The problem of determining or estimating $f(n, r)$ - the maximum possible cardinality of an r -superimposed family of subsets of $[n]$ has been considered in various papers [9, 10, 11, 12, 13, 14]. This problem can be posed as a group testing problem, which is motivated by practical problems in molecular biology. Examples include the quality control of DNA chips [6], closing the remaining gaps in the genome at the end of a sequencing project [4], [3] and clone library screening [5, 7]. As shown in [9, 14, 11],

$$\frac{c_1}{r^2} \leq \frac{\log f(n, r)}{n} \leq \frac{c_2 \log r}{r^2}, \quad (1)$$

where $c_1, c_2 > 0$ are absolute constants.

A weaker requirement is that, given the union of up to r sets, one will be able to identify at least one of those sets. Such families are called r -single-user tracing superimposed (r -SUT).

^{*}Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University. Email: nogaa@post.tau.ac.il.

[†]School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: veraa@post.tau.ac.il.

Definition 1 A family \mathcal{F} of subsets of $[n]$ is called r -single-user tracing superimposed (r -SUT) if for all choices of $\mathcal{F}_1, \dots, \mathcal{F}_k \subseteq \mathcal{F}$ with $1 \leq |\mathcal{F}_i| \leq r$,

$$\bigcup_{A \in \mathcal{F}_1} A = \bigcup_{A \in \mathcal{F}_2} A = \dots = \bigcup_{A \in \mathcal{F}_k} A$$

implies

$$\bigcap_{i=1}^k \mathcal{F}_i \neq \emptyset.$$

This problem is also motivated by applications in molecular biology, where, for example, a group of DNA sequences that carry relevant genomic information is under study, and the objective is to find at least one sequence with this information. In this problem a set of m sequences is given, containing at most r positives, i.e. sequences with the required information. In each experiment, a subset of the sequences is tested, and the outcome of the experiment is positive if and only if the subset contains at least one positive sequence. The objective is to find at least one positive sequence by conducting as few experiments as possible, where all the experiments are carried out simultaneously. In such experiments, we can view each of the given DNA sequences as a subset of the experiments (the experiments in which it participates), and in order to be able to identify at least one sequence with the relevant information, we need these subsets to form an r -SUT family. See [7] for further discussion of such applications.

Let $g(n, r)$ denote the maximum possible size of an r -SUT family of subsets of $[n]$. Thus, $m = g(n, r)$ is the maximum number of DNA sequences for which we can solve the above problem by n experiments. Csürös and Ruszinkó [7] proved that there exist constants $c_1, c_2 > 0$ such that

$$\frac{c_1}{r^2} \leq \frac{\log g(n, r)}{n} \leq \frac{c_2}{r}.$$

In this work we show that their upper bound is tight up to a constant factor. We first prove, in section 2, a lower bound of $\Omega(\frac{1}{r})$ using a probabilistic argument, and then give, in section 3, an explicit construction for fixed r , that is, an algorithm that constructs a family of the required size in time $m^{O(r)}$. Note that our lower bound, together with (1), imply that, for large r , $g(n, r)$ is much bigger than $f(n, r)$, showing that the problem of identifying one positive DNA sequence can be solved by significantly less experiments than the problem of identifying all positives.

2 The Lower Bound

Our main result is the following theorem that provides a lower bound on the maximum possible size of an r -SUT family of subsets of $[n]$. This bound determines $\log g(n, r)$ up to a constant factor. In what follows we make no attempts to optimize this absolute constant. To simplify the presentation we omit all floor and ceiling signs, whenever these are not crucial.

Theorem 1 For any $r \geq 2$ and $n \geq 20r$, there exists an r -SUT family of subsets of $[n]$ of size at least $2^{\frac{n}{20r}}$.

Fix $r \geq 2$ and $n \geq 20r$. Let $m = 2^{\frac{n}{20r}}$, and let $p = \frac{1}{r}$. Choose a family $\mathcal{F} = \{F_1, F_2, \dots, F_m\}$ of subsets of $[n]$ at random, where the subsets F_i are chosen independently as follows. Every $x \in [n]$ is chosen to be in F_i independently with probability p . We next show that with positive probability the family \mathcal{F} is r -SUT. Thus, we have to show that, with positive probability, for all choices of $\mathcal{F}_1, \dots, \mathcal{F}_k \subseteq \mathcal{F}$ such that $1 \leq |\mathcal{F}_i| \leq r$ for all $1 \leq i \leq k$ and $\cap_{i=1}^k \mathcal{F}_i = \emptyset$, the unions $\cup_{A \in \mathcal{F}_i} A$ for $1 \leq i \leq k$ are not all equal. We consider two different cases, according to the size of $\cup_{i=1}^k \mathcal{F}_i$. Proposition 2 deals with the case $|\cup_{i=1}^k \mathcal{F}_i| < 2r$, Proposition 3 with $|\cup_{i=1}^k \mathcal{F}_i| \geq 2r$, and in Proposition 4 we combine the above to complete the proof of Theorem 1.

Proposition 2 *The following holds with probability greater than $\frac{1}{2}$. For all $s < 2r$, and for all distinct $A_1, \dots, A_s \in \mathcal{F}$, there exists an element $x \in [n]$ that belongs to exactly one of the sets A_i , $1 \leq i \leq s$.*

Proof: Fix $s < 2r$ and distinct $A_1, \dots, A_s \in \mathcal{F}$. The probability that there is no element $x \in [n]$ that belongs to exactly one of these sets is

$$\begin{aligned} \left[1 - sp(1-p)^{s-1}\right]^n &\leq \left[1 - \frac{s}{r}\left(1 - \frac{1}{r}\right)^{2r-2}\right]^n \\ &\leq \left(1 - \frac{s}{r}e^{-2}\right)^n \\ &\leq e^{-e^{-2}\frac{sn}{r}} \\ &< 2^{-0.15\frac{sn}{r}}. \end{aligned}$$

Thus, the expected number of choices of $A_1, \dots, A_s \in \mathcal{F}$, $1 \leq s < 2r$, for which there is no element $x \in [n]$ that belongs to exactly one of the sets A_i , $1 \leq i \leq s$, is at most

$$\begin{aligned} \sum_{s=1}^{2r-1} \binom{m}{s} 2^{-0.15\frac{sn}{r}} &\leq \sum_{s=1}^{2r-1} m^s 2^{-0.15\frac{sn}{r}} \\ &= \sum_{s=1}^{2r-1} 2^{\frac{n}{20r}s} 2^{-0.15\frac{sn}{r}} \\ &= \sum_{s=1}^{2r-1} 2^{-0.1\frac{sn}{r}} \\ &\leq \sum_{s=1}^{2r-1} 2^{-2s} \\ &< \frac{1}{2}. \end{aligned}$$

Therefore, by Markov's inequality, the probability that there is no choice of $A_1, \dots, A_s \in \mathcal{F}$ as above is greater than $\frac{1}{2}$. \square

Proposition 3 *The following holds with probability greater than $\frac{1}{2}$. For all $t \leq r$, and for all distinct $A_1, \dots, A_r, B_1, \dots, B_t \in \mathcal{F}$,*

$$\bigcup_{i=1}^r A_i \not\subseteq \bigcup_{i=1}^t B_i.$$

Proof: Fix distinct $A_1, \dots, A_r, B_1, \dots, B_t \in \mathcal{F}$, $t \leq r$. For $x \in [n]$, the probability that $x \in \bigcup_{i=1}^r A_i$ and $x \notin \bigcup_{i=1}^t B_i$ is

$$\begin{aligned} [1 - (1-p)^r](1-p)^t &\geq \left[1 - \left(1 - \frac{1}{r}\right)^r\right] \left(1 - \frac{1}{r}\right)^r \\ &\geq \frac{1}{2}e^{-1}(1 - e^{-1}) > 0.1 \end{aligned}$$

Therefore,

$$\Pr\left(\bigcup_{i=1}^r A_i \subseteq \bigcup_{i=1}^t B_i\right) < 0.9^n,$$

and hence the expected number of choices of distinct $A_1, \dots, A_r, B_1, \dots, B_t \in \mathcal{F}$, $t \leq r$, such that

$$\bigcup_{i=1}^r A_i \subseteq \bigcup_{i=1}^t B_i.$$

is at most

$$m^{2r}0.9^n = 2^{\frac{n}{10}}0.9^n < \frac{1}{2},$$

for $n \geq 20r \geq 40$. Therefore, by Markov's inequality, the probability that there is no choice of $A_1, \dots, A_r, B_1, \dots, B_t \in \mathcal{F}$ as above is greater than $\frac{1}{2}$. \square

Proposition 4 *Any family that satisfies the properties in Propositions 2 and 3 is r -SUT. Therefore, with positive probability, the random family \mathcal{F} is r -SUT, and hence*

$$\frac{\log g(n, r)}{n} \geq \frac{1}{20r}.$$

Proof: Suppose \mathcal{F} satisfies the properties in Propositions 2 and 3. We have to show that for all $\mathcal{F}_1, \dots, \mathcal{F}_k \subseteq \mathcal{F}$ such that $1 \leq |\mathcal{F}_i| \leq r$ for all $1 \leq i \leq k$ and $\bigcap_{i=1}^k \mathcal{F}_i = \emptyset$, the unions $\bigcup_{A \in \mathcal{F}_i} A$ for $1 \leq i \leq k$ are not all equal. Consider first all such $\mathcal{F}_1, \dots, \mathcal{F}_k$ for which

$$\left|\bigcup_{i=1}^k \mathcal{F}_i\right| < 2r.$$

Since \mathcal{F} satisfies the property in Proposition 2, for all such choices of $\mathcal{F}_1, \dots, \mathcal{F}_k$ there is an element $x \in [n]$ that belongs to exactly one set B in $\bigcup_{i=1}^k \mathcal{F}_i$. Since $\bigcap_{i=1}^k \mathcal{F}_i = \emptyset$, there exists $1 \leq j \leq k$ such that $B \notin \mathcal{F}_j$, and hence $x \notin \bigcup_{A \in \mathcal{F}_j} A$. On the other hand, there is some $1 \leq j' \leq k$ for which $B \in \mathcal{F}_{j'}$. Therefore, $x \in \bigcup_{A \in \mathcal{F}_{j'}} A$, and hence $\bigcup_{A \in \mathcal{F}_j} A \neq \bigcup_{A \in \mathcal{F}_{j'}} A$, as needed.

Now consider the choices of $\mathcal{F}_1, \dots, \mathcal{F}_k$ for which

$$\left| \bigcup_{i=1}^k \mathcal{F}_i \right| \geq 2r.$$

Let $\mathcal{F}_1 = \{B_1, \dots, B_t\}$. Since $t \leq r$, $|\bigcup_{i=1}^k \mathcal{F}_i \setminus \mathcal{F}_1| \geq r$. Let A_1, \dots, A_r be distinct sets in $\bigcup_{i=1}^k \mathcal{F}_i \setminus \mathcal{F}_1$. If all the unions $\bigcup_{A \in \mathcal{F}_i} A$ for $1 \leq i \leq k$ were equal, then we would have

$$\bigcup_{i=1}^r A_i \subseteq \bigcup_{i=1}^t B_i.$$

But as \mathcal{F} satisfies the property in Proposition 3, the above does not hold for any choice of distinct $A_1, \dots, A_r, B_1, \dots, B_t \in \mathcal{F}$. Thus, no choice of $\mathcal{F}_1, \dots, \mathcal{F}_k$ with $|\bigcup_{i=1}^k \mathcal{F}_i| \geq 2r$ violates the desired property.

The assertion in each of the two propositions holds with probability exceeding $1/2$, hence they hold simultaneously with positive probability. This completes the proof of Theorem 1. \square

3 Explicit Construction

In this section we describe an explicit construction of an r -SUT family of subsets of $[n]$, of size $m = 2^{\frac{n}{20r}}$. Our construction takes time $m^{O(r)}$, and is based on a derandomization of the proof in section 2, in the spirit of the results in [1, 2], that combines the method of conditional expectations with the known constructions of small sample spaces supporting $2r$ -wise independent random variables.

Let $A = (a_{ij})$ be an $n \times m$ matrix, whose columns are the incidence vectors of a family \mathcal{F} of subsets of $[n]$. By the proofs in section 2, in order to ensure that the columns of A form an r -SUT family \mathcal{F} , it suffices to make sure that the following two conditions hold:

1. For all $s < 2r$ and all subsets $J \subseteq [m]$ of size s , there exists an index $1 \leq i \leq n$ such that there is exactly one index $j \in J$ for which $a_{ij} = 1$.
2. For all disjoint subsets $J_1, J_2 \subseteq [m]$ of size r each, there exists an index $1 \leq i \leq n$ such that for all $j \in J_1$ $a_{ij} = 0$, and there is at least one $j \in J_2$ for which $a_{ij} = 1$.

Consider the distribution $D : \{0, 1\}^m \rightarrow [0, 1]$ in which a vector $x = (x_1, \dots, x_m) \in \{0, 1\}^m$ is chosen as follows. For all $1 \leq i \leq m$, independently, let x_i be 1 with probability $\frac{1}{r}$. For any subset $J \subseteq [m]$ of size s , the probability that a vector x drawn randomly from D satisfies condition 1 with respect to J is $p_s = \frac{s}{r} (1 - \frac{1}{r})^{s-1} \geq \frac{s}{r} (1 - \frac{1}{r})^{2r-2}$. For any disjoint subsets $J_1, J_2 \subseteq [m]$ of size r each, the probability that a vector x drawn randomly from D satisfies condition 2 with respect to J_1 and J_2 is $q = (1 - \frac{1}{r})^r (1 - (1 - \frac{1}{r})^r)$. Note that this holds even if the x_i 's are only $2r$ -wise independent, instead of mutually independent.

In [1] it is shown that there is an explicit construction of a distribution $D' : \{0, 1\}^m \rightarrow [0, 1]$ with the following properties:

- D' is $2r$ -wise independent.
- $|Supp(D')| = m^{O(r)}$, and this support can be constructed in time $m^{O(r)}$.
- For all $1 \leq i \leq m$, $|Pr_{x \sim D}(x_i = 1) - Pr_{x \sim D'}(x_i = 1)| \leq \frac{1}{2m}$.

Therefore, if m is sufficiently large, we have $p'_s > 0.1\frac{s}{r}$ and $q' > 0.1$, where p'_s and q' are the probabilities corresponding to p_s and q , when choosing x according to D' instead of D .

We now use the greedy algorithm to construct the matrix A . The rows of A are chosen one by one from the support of the distribution D' as follows. Let A_t denote the $t \times m$ matrix obtained by the first t iterations. For $0 \leq t \leq n$, let X_t^s , $1 \leq s < 2r$, denote the number of subsets $J \subseteq [m]$ of size s for which there is no $i \in [t]$ that satisfies condition 1, and let Y_t denote the number of pairs of disjoint subsets $J_1, J_2 \subseteq [m]$ of size r each, for which there is no $i \in [t]$ that satisfies condition 2. Define

$$W_t = \sum_{s=1}^{2r-1} X_t^s (1 - p'_s)^{n-t} + Y_t (1 - q')^{n-t}.$$

After choosing the first $t - 1$ rows of A , the greedy algorithm chooses, in iteration number t , to add to A the vector in the support of D' that minimizes W_t . Note that

$$\begin{aligned} W_0 &= \sum_{s=1}^{2r-1} X_0^s (1 - p'_s)^n + Y_0 (1 - q')^n \\ &\leq \sum_{s=1}^{2r-1} m^s (1 - 0.1\frac{s}{r})^n + m^{2r} (1 - 0.1)^n \\ &= \sum_{s=1}^{2r-1} 2^{\frac{sn}{20r}} e^{-\frac{0.1sn}{r}} + 2^{\frac{n}{10}} 0.9^n \\ &< 1, \end{aligned}$$

as shown in section 2.

We claim that, for all $1 \leq t \leq n$, the expected value of W_t , given the outcome of the first $t - 1$ iterations, is at most W_{t-1} . Indeed,

$$\begin{aligned} E[W_t | X_{t-1}^1, \dots, X_{t-1}^{2r-1}, Y_{t-1}] &= \sum_{s=1}^{2r-1} E[X_t^s | X_{t-1}^s] (1 - p'_s)^{n-t} + E[Y_t | Y_{t-1}] (1 - q')^{n-t} \\ &\leq \sum_{s=1}^{2r-1} X_{t-1}^s (1 - p'_s)^{n-t+1} + Y_{t-1} (1 - q')^{n-t+1} \\ &= W_{t-1}. \end{aligned}$$

Therefore, for all t , the greedy algorithm chooses a vector so that $W_t \leq W_{t-1}$, and hence at the end $W_n \leq W_0 < 1$. But

$$W_n = \sum_{s=1}^{2r-1} X_n^s + Y_n.$$

Therefore, $X_n^1 = \dots = X_n^{2r-1} = Y_n = 0$, that is, the matrix $A = A_n$ satisfies conditions 1 and 2, and thus its columns are the incidence vectors of an r -SUT family of subsets of $[n]$. Obviously, the above process can be performed in time $O(m^{2r} \cdot |\text{Supp}(D')| \cdot n) = m^{O(r)}$.

4 Concluding Remarks

It may be interesting to find explicit constructions that provide r -SUT families of size $m = 2^{\Omega(\frac{n}{r})}$ in time $m^{O(1)}$ even when r is not a constant.

Csűrös and Ruszinkó [7] gave an upper bound of $2^{\frac{2n}{r}}$ for $g(n, r)$, which is the maximum size of an r -SUT family, and we proved a lower bound of $2^{\frac{n}{20r}}$. Although the constant $\frac{1}{20}$ in our proof can be easily improved, the problem of finding the exact constant remains open (and seems difficult).

Finding the exact constant may be interesting even for small values of r . For the first interesting case, $r = 2$, a result of Coppersmith and Shearer [8] implies that $g(n, 2) \leq 2^{(\frac{1}{2} + o(1))n}$, whereas the simple probabilistic argument given in Proposition 5 shows that there is a 2-SUT family of subsets of $[n]$ of size $\Omega(2^{\frac{n}{3}})$, that is, $g(n, 2) \geq \Omega(2^{\frac{n}{3}})$.

Proposition 5 *For all n , there exists a 2-SUT family of subsets of $[n]$ of size $\Omega(2^{\frac{n}{3}})$.*

Proof: Let $m = \frac{1}{2} \cdot 2^{\frac{n}{3}}$, and put $p = 1 - \frac{\sqrt{2}}{2}$. Choose a family $\mathcal{F} = \{F_1, F_2, \dots, F_m\}$ of subsets of $[n]$ at random, where the subsets F_i are chosen independently as follows. Every $x \in [n]$ is chosen to be in F_i independently with probability p .

A family \mathcal{G} is 2-SUT if the following two conditions hold:

- For all distinct $A, B, C, D \in \mathcal{G}$, $A \cup B \neq C \cup D$.
- For all distinct $A, B, C \in \mathcal{G}$, the unions $A \cup B$, $B \cup C$ and $A \cup C$ are not all equal.

For each fixed distinct $A, B, C, D \in \mathcal{F}$,

$$\begin{aligned} Pr(A \cup B = C \cup D) &= \left[(1-p)^4 + (1 - (1-p)^2)^2 \right]^n \\ &= \left[\left(\frac{\sqrt{2}}{2} \right)^4 + \left(1 - \left(\frac{\sqrt{2}}{2} \right)^2 \right)^2 \right]^n \\ &= 2^{-n}. \end{aligned}$$

Therefore, the expected number of distinct $A, B, C, D \in \mathcal{F}$ for which $A \cup B = C \cup D$ is at most

$$m^4 2^{-n} = \frac{1}{16} \cdot 2^{\frac{4n}{3}} \cdot 2^{-n} = \frac{1}{16} \cdot 2^{\frac{n}{3}} = \frac{m}{8}.$$

For distinct $A, B, C \in \mathcal{F}$,

$$\begin{aligned} \Pr(A \cup B = B \cup C = A \cup C) &= \left[1 - 3p(1-p)^2\right]^n \\ &= \left[1 - 3\left(1 - \frac{\sqrt{2}}{2}\right)\left(\frac{\sqrt{2}}{2}\right)^2\right]^n \\ &< 2^{-\frac{4}{5}n}. \end{aligned}$$

Therefore, the expected number of distinct $A, B, C \in \mathcal{F}$ for which $A \cup B = B \cup C = A \cup C$ is at most

$$m^3 2^{-\frac{4}{5}n} = \frac{1}{8} \cdot 2^n \cdot 2^{-\frac{4}{5}n} = \frac{1}{8} \cdot 2^{\frac{n}{5}} < \frac{m}{8},$$

with room to spare.

Therefore, there is a family \mathcal{F} with at most $\frac{m}{4}$ forbidden configurations of the two kinds. By removing one set from each forbidden configuration, we are left with a 2-SUT family (possibly with a few identical members) of size at least $\frac{3}{4}m = \Omega(2^{\frac{n}{3}})$, and this family cannot contain three identical members. Thus, $g(n, 2) \geq \Omega(2^{\frac{n}{3}})$. \square

References

- [1] N. Alon, L. Babai and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem, *J. Algorithms* 7 (1986), 567-583.
- [2] N. Alon, D. Moshkovitz and M. Safra, Algorithmic construction of sets for k -restrictions, submitted.
- [3] N. Alon, R. Beigel, S. Kasif, S. Rudich and B. Sudakov, Learning a hidden matching, *Proc. of the 43th IEEE FOCS, IEEE(2002)*, 197-206. Also: *SIAM J. Computing* 33 (2004), 487-501.
- [4] R. Beigel, N. Alon, M. S. Apaydin, L. Fortnow and S. Kasif, An optimal procedure for gap closing in whole genome shotgun sequencing, *Proc. 2001 RECOMB, ACM Press*, 22-30.
- [5] D. J. Balding, W. J. Bruno, E. Knill and D. C. Torney, A comparative survey of non-adaptive pooling designs, *Genetic mapping and DNA sequencing, IMA Volumes in Mathematics and its Applications*, T. P. Speed and M. S. Waterman eds, New York: Springer-Verlag, vol. 81, 1996, 133-154.
- [6] C. C. Colbourn, A. C. H. Ling and M. Tompa, Construction of optimal quality control for oligo arrays, *Bioinformatics* 18(4), 529-535 (2002).

- [7] M. Csűrös and M. Ruszinkó, Single-user tracing and disjointly superimposed codes, to appear.
- [8] D. Coppersmith and J. B. Shearer, New bounds for union-free families of sets, *The Electronic Journal of Combinatorics*, 5(1), 1998, #R39.
- [9] A. G. Dyachkov and V. V. Rykov, Bounds on the Length of Disjunctive Codes, *Problemy Peredachi Informatsii*, Vol. 18, no. 3, 158-166 (1982).
- [10] P. Erdős, P. Frankl, and Z. Füredi, Families of finite sets in which no set is covered by the union of r others, *Israel J. Math.*, vol. 51, 79-89 (1985).
- [11] Z. Füredi, A note on r -cover-free families, *Journal of Combinatorial Theory Series A*, vol. 73, no. 1, 172-173 (1996).
- [12] F. K. Hwang, V. T. Sós, Non-adaptive hypergeometric group testing, *Stud. Sci. Math. Hungarica*, vol. 22, 257-263 (1987).
- [13] W. H. Kautz and R. C. Singleton, Nonrandom binary superimposed codes, *IEEE Transactions on Information Theory*, vol. 10, 363-377 (1964).
- [14] M. Ruszinkó, On the Upper Bound of the size of the r -cover-free families, *Journal of Combinatorial Theory Series A*, vol. 66, no. 2, 302-310 (1994).