Fifteenth International Workshop on Algebraic and Combinatorial Coding Theory
June 18-24, 2016, Albena, Bulgaria *pp. 133–138*

Finding one of D defective elements in the additive group testing model [1]

C. Deppe cdeppe@math.uni-bielefeld.de
Department of Mathematics, University of Bielefeld, Germany
V. Lebedev lebedev37@yandex.ru
Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Russia

Abstract. In contrast to the classical goal of group testing a searcher want to find one defective element among D defective elements. We analyse the additive group testing model. We construct an adaptive testing algorithm and show that the algorithm is optimal.

1 Introduction

Before we introduce the problem, let us recall some definitions and notations. In combinatorial group testing, we are given a set of N elements, D of them are defective. A searcher is interested in identifying all or some defectives. Group testing is of interest for many applications like in multiaccess communication, coding, library screening and molecular biology. For an overview of results and applications we refer to the books [1], [2] and [3].

Formally, we have a set of elements denoted by $[N] := \{1, 2, \ldots, N\}$ and a set of defective elements $\mathcal{D} \subset [N]$. We denoted by $D = |\mathcal{D}|$ its cardinality. Throughout the paper we assume that a searcher knows D.

We denote by $[i, j]$ the set of integers $\{x \in [N] : i \leq x \leq j\}$ and by $2^{[N]}$ the set of all subsets of $[N]$.

The aim of a searcher is to determine a goal set $\mathcal{G} \subset [N]$ with some properties, for example in classical group testing $\mathcal{G} = \mathcal{D}$. To determine $\mathcal{G}$ the searcher can choose sets (questions) $S_i \subset [N]$ for $1 \leq i \leq n$ and asks for the values $t(S_i)$ (answers) of a test function $t : 2^{[N]} \to \mathbb{R}$.

Definition 1 *Let t be a test function, $s = (S_1, S_2, \ldots, S_n)$ be a sequence of sets $S_i \subset [N]$, and $t(s) := (t(S_1), \ldots, t(S_n))$. We call $(s, t(s), n)$ a test with test length n, if the searcher uniquely determine $\mathcal{G}$.*

[1] This research is partially supported by the Russian Foundation for Basic Research, project no 15-01-08051.

We consider a subset $\mathcal{S} \subset [N]$. For the classical group testing model a test function $t : 2^{[N]} \to \{0,1\}$ is defined by

$$t^{(Cla)}(S) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| = 0 \\ 1 & , \text{ otherwise.} \end{cases} \tag{1}$$

We consider also the threshold group testing function without gap.

$$t^{(Thr)}(S) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < u \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq u. \end{cases} \tag{2}$$

This kind of test function was introduced in [4] and some results we give in [5].

For the additive model we have the test function

$$t^{(Add)}(S) = |\mathcal{S} \cap \mathcal{D}| \tag{3}$$

We distinguish between adaptive and nonadaptive tests. We call a test nonadaptive if all questions are specified simultaneously.
A test is called adaptive if all questions are conducted one by one, and outcomes $(t(S_1), \ldots, t(S_{i-1}))$ of previous questions are known at the time of determining the current question S_i. We consider only adaptive tests.

Problem 1 *Let us call a test successfull if for any $\mathcal{D}$ we have $G = \{i\}$, $i \in \mathcal{D}$. We assume that D and n are given. How big can we choose N to ensure a successful test?*

Problem 2 *Let us fix some $j \in [1, d]$ and call a test successfull if for any $\mathcal{D}$ we have $G = \{d_j\}$, where $\mathcal{D} = \{d_1, \ldots, d_D\} \subset [N]$ and $d_1 < d_2, \cdots < d_D$. We assume again that D and n are given. How big can we choose N in this case to ensure a successful test?*

Denote by $N_{(Thr)}(n, D, u, m)$ the maximal number of elements in a set $[N]$ such that the searcher can find m defective elements in $\mathcal{D} \subset [N]$ with the test function $t^{(Thr)}$ and test length n. In [6] we proved the following

Theorem 1 *If $D \geq u$ then $N_{(Thr)}(n, D, u, 1) = 2^n + D - 1$.*

2 Additive testing

Now we consider only additive group testing. Recall that $t^{(Add)}(S) = |\mathcal{S} \cap \mathcal{D}|$. For the proof of the next theorem we need the following notations. Each set $\mathcal{S}_i \subset [N]$ can be uniquely represented by a sequence

$$\mathbf{a}_i = (a_i(1), a_i(2), \dots, a_i(N)) \in \{0,1\}^N,$$

where $a_i(j) = 1$ if $j \in S_i$ and $a_i(j) = 0$ otherwise. Therefore we can present the sequence of questions as a matrix $(a_i(j))_{i=1,\dots,n;\ j=1,\dots,N}$ with n rows and N columns. Notice that in adaptive testing the ith row depend on the $i-1$ previous rows and the given answer. Let $h_1, h_2, \dots h_s \in \{0,1\}$. We set

$$T_{h_1 h_2 \dots h_s} := \{1 \le j \le N : \forall i \in [1,s]\ a_i(j) = h_i\} \tag{4}$$

Denote by $N_{(Add)}(n, D)$ the maximal number of elements such that the searcher can find one defective element (construct a successfull test for the Problem 1) with test length n.

Theorem 2 *We have* $N_{(Add)}(n, D) = 2^n + D - 1$.

The proof is given in Section 3. We ilustrate its idea with the following case with two defectives (see [7] for more results with two defectives in group testing).

Example Let us assume that a searcher knows that there is one defective in a set with cardinality $2^{r-1}+1$ and there is one defective another set with the same cardinality for any fixed natural r. Then the searcher needs at least r questions for finding one defective element.

Assume the sets are $[1, 2^{r-1}+1]$ and $[2^{r-1}+2, 2^r+2]$.
Indeed for $r = 1$ we have sets $[1,2]$ and $[3,4]$ and we don't know 1 or 2 and 3 or 4 is defective. So the searcher needs at least one question.
Now for any natural r denote $T_0 = [1, 2^{r-1}+1]$, $T_1 = [2^{r-1}+2, 2^r+2]$ and consider arbitrary question S, $S \subseteq [1, 2^r+2]$. We have

$T_{01} = S \bigcap T_0$, $T_{00} = T_0 \backslash T_{01}$, $T_{11} = S \bigcap T_1$, $T_{10} = T_1 \backslash T_{11}$.

We assume that a genius gives the searcher the information how many defectives are in the sets T_{01}, T_{00}, T_{11}, T_{10}. We set

$$A = \begin{cases} T_{00} & , \text{ if } |T_{01}| \le |T_{00}| \\ T_{01} & , \text{ if } |T_{00}| < |T_{01}|. \end{cases} \tag{5}$$

and

$$B = \begin{cases} T_{10} & , \text{ if } |T_{11}| \le |T_{10}| \\ T_{11} & , \text{ if } |T_{10}| < |T_{11}|. \end{cases} \tag{6}$$

For any question S it is possible to get an answer, such that there is one defective in the set A and there is one defective in the set B.
Therefore $|A| \geq 2^{r-2}+1$ and $|B| \geq 2^{r-2}+1$. Thus by induction the assumption in the example is correct.

Denote by $N_{(Add)}(n, D, j)$ the maximal number of elements such that the searcher can find the jth defective element (construct a successfull test for the Problem 2) with test length n.

Theorem 3 *We have* $N_{(Add)}(n, D, j) = 2^n + D - 1$ *for* $1 \leq j \leq D$.

Proof
It holds $N_{(Add)}(n, D, j) \leq N_{(Add)}(n, D) = 2^n + D - 1$. Therefore we have to give a successful test for $N = 2^n + D - 1$. Recall that d_j denotes the jth defective.

Our assumption is that for any set $\mathcal{T}$ with $|\mathcal{T}| \leq 2^n + x - 1$ containing x defectives the searcher can find d_z with n questions for any $1 \leq z \leq x$.

For $n = 1$ the searcher sets $S_1 = [1, z]$.
The answer could be

$$\begin{cases} z - 1 & \textit{then } d_z = z + 1 \\ z & \textit{then } d_z = z \end{cases} .$$

Therefore the assumption is true for $n = 1$.

Let us presume that for $n - 1$ the assumption is true.
For $[N] = [1, 2^n + D - 1]$ the searcher sets $S_1 = [1, 2^{n-1} + j - 1]$
The answer could be

k≥j In this case $\mathcal{T} = S_1$, $x = k$ and $z = j$.

Therefore $|\mathcal{T}| = 2^{n-1} + j - 1 \leq 2^{n-1} + k - 1$ and by induction we have a test with $n - 1$ questions.

k<j In this case $\mathcal{T} = [N] \backslash S_1$, $x = D - k$ and $z = j - k$.

Because of $\mathcal{T} = [2^{n-1} + j, 2^n + D - 1]$ we have

$|\mathcal{T}| = 2^{n-1} + D - j \leq 2^{n-1} + D - k - 1$ and by induction we have a test with $n - 1$ questions.

From the assumption follows the Theorem 3. □

3 Proof of Theorem 2

It is obviously that $N_{(Add)}(n, D, 1) \geq N_{(Thr)}(n, D, 1, 1) = 2^n + D - 1$.
Assume that $N = 2^n + D$. We will show that the searcher cannot find one defective element with n questions.
After r questions we have a partition of N elements to 2^r sets

$$T_{h_1 h_2 \ldots h_r},\ h_i \in \{0, 1\}$$

as defined above.
We assume that some genius tell the searcher how many defectives are in any set $T_{h_1 h_2 \ldots h_r}$ and denote it by $D_{h_1 h_2 \ldots h_r}$.
This give the searcher more information as he has in the additive model.

We will show that
If every set $T_{h_1 h_2 \ldots h_s}$ containing one or more defectives has cardinality

$$|T_{h_1 h_2 \ldots h_s}| = 2^{n-s} + D_{h_1 h_2 \ldots h_s}$$

then the searcher needs more than $n-s$ questions to find one defective element.
We proof this by induction in $n - s$.

$n - s = 1$: It is easy to see that for this case $|T_{h_1 h_2 \ldots h_s}| = 2 + D_{h_1 h_2 \ldots h_s}$ and it is not possible in the worst case to identify one defective with one question.

The question S_s partition each set $T_{h_1 h_2 \ldots h_{s-1}}$ which contains defectives to the sets $T_{h_1 h_2 \ldots h_{s-1} 0}$ and $T_{h_1 h_2 \ldots h_{s-1} 1}$. It is possible that

$$t_{(Add)}(T_{h_1 h_2 \ldots h_s}) = \begin{cases} 0 & , if\ |T_{h_1 h_2 \ldots h_s}| \leq 2^{n-s} \\ D_{h_1 h_2 \ldots h_{s-1}} & , if\ |T_{h_1 h_2 \ldots h_s}| \geq 2^{n-s} + D_{h_1 h_2 \ldots h_{s-1}} \\ |T_{h_1 h_2 \ldots h_s}| - 2^{n-s} & , otherwise \end{cases}$$

(We use that $D_{h_1 h_2 \ldots h_{s-1} 1} + D_{h_1 h_2 \ldots h_{s-1} 0} = D_{h_1 h_2 \ldots h_s}$).

In the third case it holds $|T_{h_1 h_2 \ldots h_s}| = 2^{n-s} + D_{h_1 h_2 \ldots h_s}$ and both sets $T_{h_1 h_2 \ldots h_{s-1} 1}$ and $T_{h_1 h_2 \ldots h_{s-1} 0}$ contains defectives.

If we are in one of the first two cases.
A genius will tell the searcher some more nondefectives, such that the set without defectives increases and get the cardinality 2^{n-s} and the other set with defectives ($h_s = 0$ or $h_s = 1$) has cardinality $|T_{h_1 h_2 \ldots h_s}| = 2^{n-s} + D_{h_1 h_2 \ldots h_s}$.

Thus by induction the searcher needs more than n questions and therefore we have only a successful test, if $N \leq 2^n + D - 1$.

□

References

[1] R. Ahlswede and I. Wegener, Suchprobleme, Teubner Verlag, Stuttgart, 1979, Russian Edition: Zadatsi Poiska, MIR, 1982, English Edition: *Search Problems, Wiley-Interscience Series in Discrete Mathematics and Optimization*, 1987.

[2] D.Z. Du and F.K. Hwang , *Combinatorial Group Testing and its Applications, 2nd edition, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, Series on Applied Mathematics*, 12, 2000.

[3] D.Z. Du and F.K. Hwang , *Pooling Designs and Nonadaptive Group Testing. Important Tools for DNA Sequencing, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, Series on Applied Mathematics*, 18, 2006.

[4] P. Damaschke, Threshold group testing, General Theory of Information Transfer and Combinatorics, R. Ahlswede et al. editors, *Lecture Notes in Computer Science*, Vol. 4123, Springer Verlag, 2006, 707-718.

[5] R. Ahlswede, C. Deppe, V. Lebedev, Bounds for threshold and majority group testing, in *2011 IEEE International Symposium on Information Theory*, Sankt-Peterburg, Russia, Aug. 1-5, 2011, 69–73.

[6] R. Ahlswede, C. Deppe, and V. Lebedev, Finding one of D defective elements in some group testing models. *Probl. Inf. Transm.*, Vol. 48, No. 2, 2012, 173-181.

[7] C. Deppe and V. Lebedev, Group testing problem with two defectives, *Probl. Inf. Transm.*, Vol. 49, No. 4, 2013, 375-381.