Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory
June 15-21, 2012, Pomorie, Bulgaria *pp. 15–20*

Finding one of D defective elements in some group testing models[1]

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

Abstract. In contrast to the classical goal of group testing we want to find m defective elements among D ($m \leq D$) defective elements. We analyse two different test functions. We give adaptive strategies and lower bounds for the number of tests and show that our strategy is optimal for $m = 1$.

1 Introduction

Group testing is of interest for many applications like in molecular biology. For an overview of results and applications we refer to the books [1] and [2].
We want to find m of D defective elements. These study was motivated by [3] and [4]. We denote by $[N] := \{1, 2, \ldots, N\}$ the set of elements, by $\mathcal{D} \subset [N]$ the set of defective elements, by $D = |\mathcal{D}|$ its cardinality, and by $[i, j]$ the set of integers $\{x \in \mathcal{N} : i \leq x \leq j\}$. Throughout the paper we consider worst case analysis.

The classical group testing problem is to find the unknown subset $\mathcal{D}$ of all defective elements in $[N]$.

For a subset $\mathcal{S} \subset [N]$ a test $t_{\mathcal{S}}$ is the function $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1\}$ defined by

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

We define search strategies as in [5]. In classical group testing a strategy is called successful, if we can **uniquely determine** $\mathcal{D}$. Here we call a strategy successful if we can find one element of $\mathcal{D}$.

[1]This research is partially supported by the Russian Foundation for Basic Research, project 12-01-00905 and by the DFG, project AH46/7-1 "General Theory of Information Transfer".

Let f be a function $f : [0, N] \to \mathbb{R}^+$. We define *general group tests with density* as $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1\}$, defined by

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < f(|\mathcal{S}|) \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq f(|\mathcal{S}|). \end{cases} \quad (2)$$

In [4] the case $f(|\mathcal{S}|) = \alpha|\mathcal{S}|$ is considered. The authors assume that a lower bound of the cardinality of $\mathcal{D}$ is known. **The goal is to find** $m \leq D$ **defective elements.**

In **majority group testing** (defined in [6] and more general in [7]) we have two functions $f_1, f_2 : \{0, 1, \ldots, N\} \to \mathbb{R}^+$ which put weights on the number D of defective elements and $f_1(D) \leq f_2(D)\ \forall D \in [0, 1, \ldots, N]$.
We describe the structure of tests $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1, \{0, 1\}\}$ as follows

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < f_1(D) \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq f_2(D) \\ \{0, 1\} & , \text{ otherwise} \\ & \text{(the result can be arbitrary 0 or 1).} \end{cases} \quad (3)$$

In [7] it is assumed that the searcher does not know the cardinality of $\mathcal{D}$ but knows some upper bound. In majority group testing **it is not always possible to find the set** $\mathcal{D}$ **of all defective elements** (see [7], [8]).In general, one can **find a family** $\mathbb{F}$ **of sets, which contains** $\mathcal{D}$. This family depends on f_1 and f_2, on $\mathcal{D}$, and on the strategy used. In this case we call a strategy successful, if we can find an $\mathbb{F}$ with the smallest possible size.

Now we put the ideas of these two models together such that there are two functions $f_1, f_2 : [0, N] \times [0, N] \to \mathbb{R}^+$ with $f_1(D, S) \leq f_2(D, S)$ for all values of D and S.
We define a test $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1, \{0, 1\}\}$ as follows

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < f_1(D, |\mathcal{S}|) \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq f_2(D, |\mathcal{S}|) \\ \{0, 1\} & , \text{ otherwise} \\ & \text{(the result can be arbitrary 0 or 1).} \end{cases} \quad (4)$$

For this test function denote by $n(N, D, m)$ the minimal number of tests for finding m defective elements.
The following lower bound for the minimal number of test is a generalization of a theorem in [4]. They give this lower bound for $f_1(D, |\mathcal{S}|) = f_2(D, |\mathcal{S}|) = \alpha|\mathcal{S}|$.

Theorem 1 $n(N, D, 1) \geq \lceil \log(N - D + 1) \rceil$

Let us assume that we have a successful strategy s which finds a defective element with $n = n(N, D, 1)$ tests and $n < \lceil \log(N - D + 1) \rceil$.
Depending on the n test results we have at most 2^n different possible results for a defective element, we denote them by $\mathcal{E}$. It holds by assumption that $|\mathcal{E}| \leq 2^n < N - D + 1$. Therefore $|[N] \backslash \mathcal{E}| > D - 1$ and there exists a set $\mathcal{F} \subset [N] \backslash \mathcal{E}$ with $|\mathcal{F}| = D$. Now we consider the case $\mathcal{D} = \mathcal{F}$. It is obvious now that strategy s we cannot find any defective element with n tests.

We denote by $n_{(Cla)}(N, D, m)$ the minimal number of tests (1) of finding m defective elements.

Proposition 1 $n_{(Cla)}(N, D, 1) \leq \lceil \log(N - D + 1) \rceil$

Proposition 1 together with Theorem 1 implies the following

Corollary 1 *1.* $n_{(Cla)}(N, D, 1) = \lceil \log(N - D + 1) \rceil$,

2. $n_{(Cla)}(N, D, m) \leq m \lceil \log(N - D + 1) \rceil$.

2 Threshold test function without gap

We consider now the test function

<u>Threshold group testing without gap:</u> $f(D, |\mathcal{S}|) = u$. Thus

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < u \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq u. \end{cases} \tag{5}$$

This kind of test was introduced in [8] and called threshold group testing without gap. First we assume that we know D.

We denote by $n_{(Thr)}(N, D, u, m)$ the minimal number of tests (5) for finding m defective elements , if we have N elements with D defectives and $f(D, |\mathcal{S}|) = u$.

Our first goal is to find one defective element.

Proposition 2 *If* $D \geq u$ *then* $n_{(Thr)}(N, D, u, 1) \leq \lceil \log(N - D + 1) \rceil$, *otherwise it is not possible to find any defective element.*

We give a strategy which needs $\lceil \log(N - D + 1) \rceil$ tests. The idea of the proof is to partition the set of N elements into the subsets $\mathcal{I}_1 = [1, u - 1]$, $\mathcal{I}_2 = [u, N - D + u]$, and $\mathcal{I}_3 = [N - D + u + 1, N]$. In $\mathcal{I}_2$ there is of course at least one defective, because the union of the two other subsets has cardinality $D - 1$. We can find a defective element in $\mathcal{I}_2$ by the following strategy with $\lceil \log(N - D + 1) \rceil$ tests.

We start with the test set

$$\mathcal{S}_1 = \{1, \ldots, u-1, u, \ldots, (u-1) + \lceil \frac{m(1)}{2}(N-D+1) \rceil\},$$

where $m(1) = 1$.
Inductively, we set $m(j) = \begin{cases} 2m(j-1) - 1 & \text{if } t_{S_{j-1}}(\mathcal{D}) = 1 \\ 2m(j-1) + 1 & \text{if } t_{S_{j-1}}(\mathcal{D}) = 0, \end{cases}$
and $\mathcal{S}_j = \{1, \ldots, u-1, u, u+1, \ldots, (u-1) + \lceil \frac{m(j)}{2^j}(N-D+1) \rceil\}$.
After $\lceil \log(N-D+1) \rceil$ tests we can find an i such that $t_{[1,i]} = 1,\ t_{[1,i-1]} = 0$ because it is clear that $t_{[1,u-1]} = 0$ and $t_{[1,N-D+u]} = 1$. Thus using this strategy we find an defective element at the position i.

From Theorem 1 and Proposition 2 we get the following

Theorem 2 $n_{(Thr)}(N, D, u, 1) = \lceil \log(N-D+1) \rceil$, *if* $D \geq u$.

3 Density tests

The test model

Group testing with density tests: $f(D, |\mathcal{S}|) = \alpha|\mathcal{S}|$ for all values. Thus

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < \alpha|\mathcal{S}| \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq \alpha|\mathcal{S}|. \end{cases} \tag{6}$$

was considered in [4].

Let $n_{(Den)}(N, D, m, \alpha)$ be the minimal number of tests (6) for finding m defective elements, if we have N elements with D defectives. In [4] the authors obtain the following bounds for $n_{(Den)}(N, D, m, \alpha)$ assuming $D \geq \alpha N$

$$\lceil \log N \rceil + \max_{N' \leq 2\frac{m}{\alpha}} n_{(Den)}(N', m, m, \alpha) \geq n_{(Den)}(N, D, m, \alpha), \tag{7}$$

$$\lceil \log N \rceil \geq n_{(Den)}(N, D, 1, \alpha). \tag{8}$$

In general they show that

$$\log(N-D+1) \leq n_{(Den)}(N, D, 1, \alpha). \tag{9}$$

We will give a strategy which is optimal for $D \geq \alpha N$ (it needs $\lceil \log(N-D+1) \rceil$ questions).

Let us define

$$s_i = \lceil \frac{2^{n-i}-1}{1-\alpha} \rceil$$

where $i = 1, 2, \ldots, n-1$ and $s_n = 1$.
For given D we choose the maximal n such that

$$D > \sum_{i=1}^{n} s_i - 2^n + 1. \tag{10}$$

Theorem 3 *Let (10) be fulfilled and $N \leq 2^n + D - 1$ then after n tests of the strategy above we will find one defective element.*

Corollary 2 *If $D \geq \alpha N$ then $n_{(Den)}(N, D, 1) = \lceil \log(N - D + 1) \rceil$.*

References

[1] 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.

[2] 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.

[3] G. Katona, Finding at least one defective element in two rounds, Dagstuhl, 2009.

[4] D. Gerbner, B. Keszegh, D. Palvölgyi and G. Wiener, Search with density tests, preprint 2010.

[5] 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.

[6] V. Lebedev, Separating codes and a new combinatorial search model, *Probl. Inf. Transm.* **46**, No. 1, 1-6, 2010.

[7] R. Ahlswede, C. Deppe, and V. Lebedev, Finding one of D defective elements in some group testing models. *Probl. Inf. Transm.* to appear.

[8] 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, 707-718, 2006.

[9] F.K. Hwang, A method for detecting all defective members in a population by group testing, *J. Amer. Statist. Assoc.* **67**, 605–608, 1972.

[10] A. D'yachkov and V. Rykov, Optimal Superimposed Codes and Designs. for Renyi's Search Model, *Journal of Statistical Planning and Inference*, Vol. 100, No. 2, 281–302, 2002.

[11] R. Ahlswede, C. Deppe, V. Lebedev, Majority group testing with density tests, in *2011 IEEE International Symposium on Information Theory*, Sankt-Peterburg, Russia, Aug. 1-5, 2011, 326–330.

[12] 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.