LARGE SYSTEMS

Group Testing Problem with Two Defectives

C. Deppe\textsuperscript{a1} and V. S. Lebedev\textsuperscript{b2}

\textsuperscript{a}Department of Mathematics, University of Bielefeld, Germany
cdeppe@mathematik.uni-bielefeld.de

\textsuperscript{b}Kharkevich Institute for Information Transmission Problems,
Russian Academy of Sciences, Moscow
lebed37@iitp.ru

Received March 19, 2013; in final form, June 14, 2013

Abstract—We consider the classical $(2, N)$ group testing problem, i.e., the problem of finding two defectives among $N$ elements. We propose a new adaptive algorithm such that for $N = \left\lfloor 2^{2t+1} - t \cdot 2^t \right\rfloor$ the problem can be solved in $t$ tests.

DOI: 10.1134/S0032946013040078

1. INTRODUCTION

Many people fond of recreational mathematics get acquainted with testing problems already in their schooldays. One of such problems is the well-known problem of finding a false coin using a balance scale in the fewest number of weighings. Another such problem is as follows: an analytical balance is used, and for a group of tested coins we can determine whether the group contains a false coin or not. Search strategies can be adaptive and nonadaptive. In contrast to adaptive strategies, nonadaptive strategies are such that all tests are known beforehand and do not change depending on results of tests already made.

For the case of one false coin, solutions of such problems have been well known for a long time. Surprisingly, even for the case of two coins precise answers for most of such problems are still unknown.

Let us formulate the problem under consideration in somewhat different terms. Assume that there are $N$ users who transmit their messages in the form of binary sequences of length $t$. During transmission through a channel, exactly two users can be active (i.e., transmit their messages according to a transmission strategy designed beforehand). At each time instant (from 1 to $t$), the channel output is 0 if both users transmit zeros at this instant, and in all other cases the channel output is 1. A transmission strategy must be such that, given the channel output, one can find out which users transmitted their messages. As above, if a transmission strategy may change depending on the channel output at each time instant, the strategy is said to be adaptive. We will consider adaptive strategies only.

In the classical group testing model, there is a set $[N] := \{1, \ldots, N\}$ of elements containing a subset $\mathcal{D} \subset [N]$ of defectives. The main problem of group testing is determining $\mathcal{D}$ in the fewest number of tests. Each test is some subset of $[N]$. It is assumed that there is a test function which for any subset $\mathcal{S} \subset [N]$ indicates the presence of a defective in this subset (gives an answer to the

\textsuperscript{1} Supported in part by the German Research Council (DFG), project “General Theory of Information Transfer.”

\textsuperscript{2} Supported in part by the Russian Foundation for Basic Research, project no. 12-01-00905-a, and German Research Council (DFG), project “General Theory of Information Transfer.”
test). Formally, a test function \( f_S : 2^{|N|} \to \{0, 1\} \) can be defined as follows:

\[
f_S(D) = \begin{cases} 
0 & \text{if } |S \cap D| = 0, \\
1 & \text{if } |S \cap D| > 0.
\end{cases}
\]  

(1)

A set of tests forms a search algorithm. We say that a search algorithm is successful if after applying it we can uniquely determine \( D \) from the answers \( f_{S_1}, \ldots, f_{S_t} \). Algorithms can be adaptive and nonadaptive. In an adaptive algorithm, when choosing a test one can use results of previous tests. In a nonadaptive algorithm, all tests are independent. In this paper we consider adaptive search algorithms only.

Below we need the following notation. Let \( |D| = D \) be the number of defectives, and \( N_t(D) \) the largest number of elements among which \( D \) defectives can be found in \( t \) tests. For an adaptive algorithm \( a = a(N, D, t) \), denote by \( a_t(D) \) the maximal number of elements for which it is proved that the \((D, a_t(D))\) problem can be solved in \( t \) tests, i.e., algorithm \( a \) is successful. Thus, \( a_t(D) \) is a lower bound for \( N_t(D) \).

A description of other interesting combinatorial search and group testing models can be found in [1, 2]. Basic problems of search theory, along with sorting and identification problems, are also presented in [3].

As was noted above, it is well known that \( N_t(1) = \lceil \log_2 t \rceil \), and finding the value of \( N_t(2) \) is a hard combinatorial problem.

In [4], the authors obtained an upper bound \( N_t(2) \leq \left(2^{(t+1)/2} - 1\right) / 2 \) and proved for a search algorithm \( l \) proposed by them that \( \frac{L(2)}{N_t(2)} > 0.95 \).

This result was improved in [5], where a proposed search algorithm \( u \) yielded the bound \( \frac{u(2)}{N_t(2)} > 0.983 \). Moreover, it was also shown in [5] that for a search algorithm \( v \) there exists \( t_0 \) such that \( \frac{v(2)}{N_t(2)} > 0.995 \) for \( t \geq t_0 \). We will return to the search algorithm \( u \) in Section 4.

Let us mention some other works in this field. In [6], a search algorithm is proposed for numbers of elements that can be represented as certain sums of Fibonacci numbers. This result is worse than the above-given estimates for a large number of elements. In [7], a search algorithm \( p \) is proposed and it is shown that \( \frac{p(2)}{N_t(2)} > 0.991 \) if \( t \geq 22 \). Note also the classical paper [8], where a probabilistic setting of the problem is considered and an optimal algorithm for a special number of tests is constructed.

We emphasize once again that in this paper we only study adaptive algorithms for the case of \( D = 2 \). Throughout the paper, we analyze the worst-case scenario and want to present an algorithm which is optimal for almost all values of \( N \). More precisely, in Section 2 we describe an algorithm \( w \) such that \( \frac{w(2)}{N_t(2)} \to 1 \) as \( t \to \infty \). Section 3 is devoted to proving that \( w(2) \geq \left\lfloor 2^{t+1} - t \cdot 2^t \right\rfloor / 4 \) for \( t \) large enough \((t \geq 44)\). The fact that this inequality as well holds for \( t \leq 44 \) will be shown in Section 4.

To conclude this section, we present a result which will be used below. It was obtained in [9] for a special case where two defectives are contained in two disjoint subsets of \([N]\), one in each subset.

**Lemma.** Assume that a set \( A \subset [N] \) is known to contain exactly one defective, a set \( B \subset [N] \) is also known to contain exactly one defective, and these sets are disjoint. Then the minimal number of tests required to find the defective in \( A \), \(|A| = m\), and the defective in \( B \), \(|B| = n\), is \( \lceil \log mn \rceil \).

Other interesting results and settings of problems close to that considered here can be found in [10–12]. For a proof of the lemma, as well as references to other works on group testing, see [1].